UNIVERSITE DE FIANARANTSOA
ECOLE NATIONALE D'INFORMATIQUE

SOCIETE BLUELINE

MEMOIRE DE FIN D' ETUDES EN VUE DE L'OBTENTION DU DIPLÔME D'INGENIEUR EN INFORMATIQUE

intitulé

LES RESEAUX SAN ET NAS, SOLUTIONS DE STOCKAGE ET DE PARTAGE DE DONNEES

Présenté et soutenu par :

RAKOTOMALALANIRINA Maminiaina Christian

Encadreur professionnel : **Monsieur NANTENAINA** Tianarivo

Encadreur pédagogique : **Monsieur RABETAFIKA** Louis Haja

UNIVERSITE DE FIANARANTSOA
ECOLE NATIONALE D'INFORMATIQUE

SOCIETE BLUELINE

MEMOIRE DE FIN D' ETUDES EN VUE DE L'OBTENTION DU DIPLÔME D'INGENIEUR EN INFORMATIQUE

intitulé

LES RESEAUX SAN ET NAS, SOLUTIONS DE STOCKAGE ET DE PARTAGE DE DONNEES

Présenté et soutenu par :

RAKOTOMALALANIRINA Maminiaina Christian

Encadreur professionnel : **Monsieur NANTENAINA** Tianarivo

Encadreur pédagogique : **Monsieur RABETAFIKA** Louis Haja

REMERCIEMENTS

J’ai effectué un stage d’une durée de 5 mois au sein de la société BLUELINE. Ce fût un stage énormément enrichissant du point de vue humain et professionnel. J’ai pu grandement enrichir mes connaissances, surtout dans le domaine de l'administration systèmes et réseaux.

Je tiens à remercier toutes les personnes qui, d’une manière ou d’une autre, m’ont aidé à mener à bien ce projet.

Je tiens à remercier tout particulièrement:

- Madame Monique **RASOAZANANERA**, Présidente de l’Université de Fianarantsoa.
- Monsieur Fontaine **RAFAMANTANANTSOA**, Directeur de l'Ecole Nationale d'Informatique qui m'a donné l'autorisation d'effectuer ce stage,
- Monsieur Damien de **LAMBERTERIE**, Directeur Général de la société BLUELINE, qui a bien voulu me prendre comme stagiaire au sein de sa société,

- Monsieur Jérôme **SANTINI**, Directeur Technique au sein de la société BLUELINE, et l'ensemble de son équipe, pour leur accueil bienveillant et leurs conseils avisés, et cela malgré leur emploi du temps chargé.

- Monsieur Tianarivo **NANTENAINA**, Administrateur Systèmes et réseaux IP, mon encadreur professionnel pour toute l'aide précieuse et les nombreux conseils qu'il m'a prodigué tout au long de ce travail.

- Monsieur Louis Haja **RABETAFIKA**, mon encadreur pédagogique, pour la confiance qu'il m'a accordé en acceptant d'encadrer ce travail, pour ses multiples conseils et pour toutes les heures qu'il a consacré au relecture du document que je lui ai adressé.

- Tous les enseignants qui ont participé à ma formation durant mes études à l'Ecole Nationale d'Informatique et

- Tout le personnel de la société BLUELINE, pour son accueil chaleureux et convivial.

CURRICULUM VITAE

Maminiaina Christian **RAKOTOMALALANIRINA**
Tél. : 032 43 927 37
E-mail: hurn_08@yahoo.fr
24 ans
Célibataire

FORMATIONS ET DIPLOMES

2010-2009 : Troisième Année de Formation d'Ingénieurs à l'ENI (Ecole Nationale d'Informatique), Université de Fianarantsoa.

2009-2008 : Deuxième Année de Formation d'Ingénieurs à l'ENI, Université de Fianarantsoa.

2008-2007 : Première Année de Formation d'Ingénieurs à l'ENI, Université de Fianarantsoa.

2006-2005 : Deuxième Année de Formation de Technicien Supérieur à l'ENI, Université de Fianarantsoa.
Obtention du diplôme DUTSMI (Diplôme Universitaire de Technicien Supérieur en Maintenance des systèmes Informatique) avec Mention TRES BIEN.

2004-2005 : Première Année de Formation de Technicien Supérieur à l'ENI, Université de Fianarantsoa.

2003-2004 : Terminale (Technique) au CSFX (Collège Saint François Xavier Fianarantsoa).
Obtention du Baccalauréat Technique avec Mention BIEN.

STAGES ET EXPERIENCES PROFESSIONNELLES

Juillet. 2010 – Nov. 2010 : Stage d'une durée de 5 mois au sein de la société BLUELINE sur le thème « **Les réseaux SAN/NAS, solutions de stockage et de partage de données** ».

- Mission : Mise en place d'un système de stockage distribué.

Nov. 2009 – Jan. 2010 : Stage d'une durée de 3 mois à l'ENI sur le thème « **QOS : Quality Of Service sous OpenBSD** ».

- Mission : Création d'un logiciel pour la gestion de la bande passante.

Sep. – Nov. 2006 : Stage d'une durée de 3 mois au MAEP (Ministère de l'Agriculture de l'Elevage et de la Pêche) en vue d'obtention du DUTSMI sur le thème « **Mise en place d'un Firewall au sein du MAEP** ».

- Mission : Installation et configuration du Firewall en utilisant la distribution RedHat Enterprise Linux Server.

Sep. – Oct. 2005 : Travaux de réalisations d'une durée d'un mois à l'ENI sur le thème « Montre Electronique ».

- Mission : Création d'une montre électronique en utilisant GRAFCET.

COMPETENCES EN INFORMATIQUES

Maintenance Informatique :

- Installation matérielle et logicielle,
- Diagnostic de pannes,
- Maintenance matérielle et logicielle.

Systèmes d'Exploitation :

- **Microsoft Windows:** XP, Vista, 2003 Server, etc.
- **UNIX:** RedHat, Mandrake, Debian, OpenBSD, FreeBSD, etc.

Bureautique: Microsoft Office, OpenOffice

Méthodes d'Analyse et de Conception : MERISE, UML.

Langages de Programmation :

- Pascal, C, C++, Java, VB, .NET, Développement sous Oracle (Forms Builder).

Langages de Script : HTML, JavaScript, PHP, JSP, ASP, Python.

SG Base de données : Access, MySQL, Oracle, PostgreSQL.

Administration systèmes et réseaux :

- **Messagerie:** DNS**,** Postfix, Dovecot.
- **Sauvegarde:** Clonezila, iSCSI & OCFS (internet Small Computer System Interface & Oracle Cluster File System), NAS & SAN (Network Attached Storage & Storage Area Network), RAID.
- **Sécurité :** IPCOP, PfSense.
- **Annuaire d'authentification :** LDAP,
- **Serveur d'authentification :** FreeRadius avec MySQL, Oracle,
- **Discussion en ligne : VOIP** (Voice Over IP).

LANGUES

	Comprendre	**Parler**	**Lire**	**Rédiger**
Français	Très Bien	Très Bien	Bien	Assez Bien
Anglais	Bien	Bien	Bien	Assez Bien

SOMMAIRE

LISTE DES ABREVIATIONS

ATA	**Advanced Technology Attachment**
CIFS	**Common Internet File System**
CLVM	**Clustered Logical Volume Manager**
DAS	**Direct Attached storage**
DHCP	**Dynamic Host Configuration Protocol**
DL M	**Distributed Lock Manager**
DRBD	**Distributed Replicated Block Device**
FC	**Fibre Channel**
GFS	**Global File System**
GNBD	**Global Network Block Device**
HBA	**Host Bus Adapter**
IMAP	**Internet Message Access Protocol**
IP	**Internet Protocol**
iSCSI	**internet Small Computer System Interface**
KVM	**Kernel-based Virtual Machine**
LAN	**Local Area Network**
LVM	**Logical Volume Manager**
LUN	**Logical Unit Number**
NAS	**Network Attached Storage**
NBD	**Network Block Device**
NFS	**Network File System**
OCFS	**Oracle Cluster File System**
OS	**Operating System**
POP	**Post Office Protocol**
QoS	**Quality of Service**
RAID	**Redundant Array of Independent Disks**
SAN	**Storage Area Network**
SATA	**Serial Advanced Technology Attachment**
SCSI	**Small Computer System Interface**
SMTP	**Simple Mail Transfer Protocol**
SPOF	**Single Point of Failure**
SQL	**Structured Query Language**
TCP	**Transmission Control Protocol**

TFTP	**Trivial File Transfer Protocol**
VPN	**Virtual Private Network**
WIMAX	**Worldwide Interoperability for Microwave Access**

LISTE DES FIGURES

LISTE DES TABLEAUX

INTRODUCTION

Les volumes de données manipulées par les entreprises qui croissent de manière exponentielle provoquent une fuite en avant dans le développement des capacités de stockage afin de garantir à tout moment la disponibilité de toutes ces données.

L'approche traditionnelle du stockage par les entreprises impliquait un lien direct entre un serveur et le système de stockage local. Cette méthode de stockage par lien direct engendrent des îlots de stockage compliqués à gérer et limitent la flexibilité. Voilà pourquoi deux nouvelles approches du stockage sont apparues : le NAS et le SAN, permettant la mutualisation des données, la simplification de la gestion et l'amélioration de la flexibilité.

Le SAN ou Storage Area Network constitue une rupture technologique. Les systèmes de stockage y sont attachés par un réseau rapide entièrement dédié au stockage. C'est une solution novatrice mais coûteuse. C'est pourquoi, nombreuses entreprises se replient vers le NAS ou Network Attached Storage dans lequel l'espace de stockage est mutualisé mais où les données continuent de transiter sur le réseau local existant de l'entreprise.

La société BLUELINE, en tant que fournisseur d'accès à Internet offre de nombreuses services à ses clients: hébergement de site web, vente de nom de domaine et tant d'autres. De ce fait, elle doit impérativement assurer la disponibilité des données liées aux services offerts.

Ce projet consiste donc à effectuer une étude approfondie des deux solutions de stockage, les réseaux SAN et NAS afin de connaître celui qui répond efficacement à l'augmentation des volumes de données et aux besoins de la société. La solution retenue fera l'objet d'une analyse conceptuelle suivie d'une réalisation pratique.

Pour ce faire, ce présent mémoire est divisé en quatre parties. La première partie présente l'Ecole Nationale d'Informatique, notre école d'origine, et la société BLUELINE, notre société d'accueil. La deuxième partie effectue une analyse préalable, la description du projet et l'analyse de l'existant. La troisième partie détaille l'étude proprement dite des deux solutions afin de connaître la solution mieux adaptée et la quatrième et dernière partie est consacrée à la réalisation du projet.

Partie I

Présentation

CHAPITRE 1 :

Présentation de l'Ecole Nationale d'Informatique

CHAPITRE 2 :

Présentation de la société BLUELINE

Chapitre 1

Présentation de l'Ecole Nationale d'Informatique

1.1 Localisation et contact

L'École Nationale d'Informatique dit ENI est localisée à Tanambao Fianarantsoa, sa boîte postale porte le numéro 1487 avec le code postal 301. Son téléphone porte le numéro 75 508 01, et son adresse électronique est la suivante eni@univ-fianar.mg.

1.2 Organigramme

L'ENI est placée sous la tutelle académique et administrative de l'Université de Fianarantsoa. La figure 1 présente l'organigramme de l'Ecole.

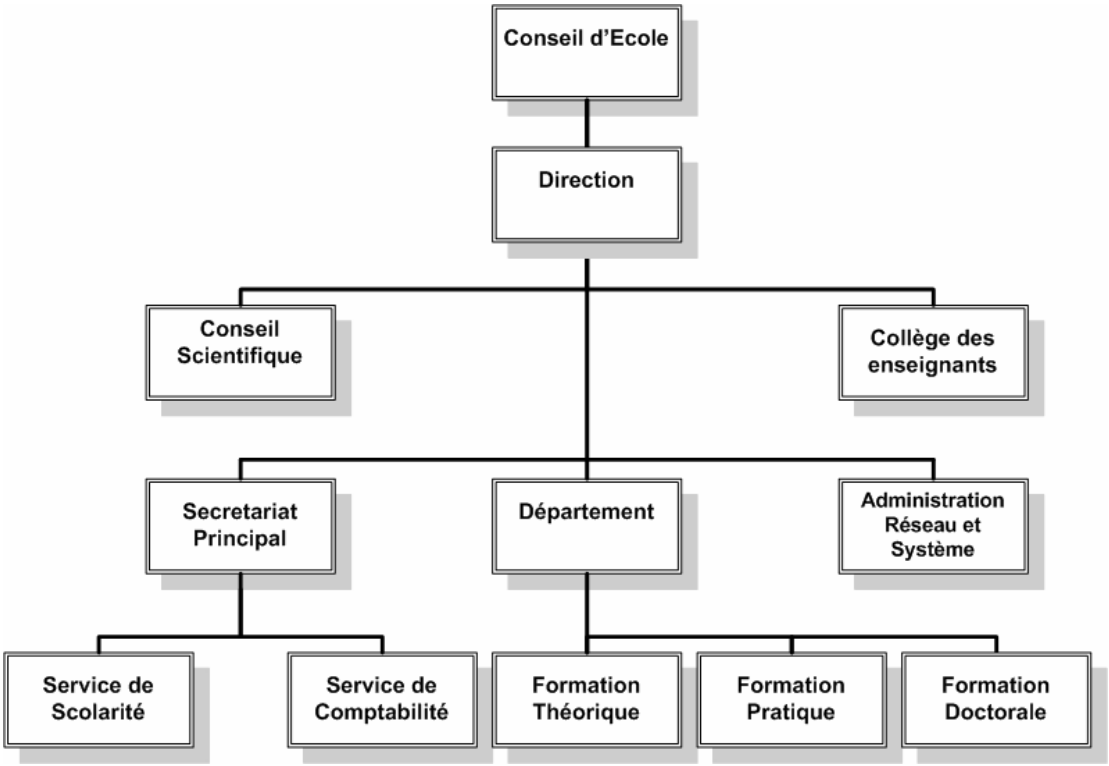

Figure 1: Organigramme de l'ENI

1.3 Missions et historiques

L'ENI de l'Université de Fianarantsoa constitue à l'heure actuelle la pépinière des élites informaticiennes malgaches. On peut considérer cette Ecole Supérieure comme la vitrine et la mesure de l'avancée technologique du Pays.

Elle se positionne dans le système socio-éducatif malgache comme le plus puissant vecteur de diffusion et de vulgarisation des connaissances et des technologies informatiques.

Créée par le Décret N° 83-185 du 24 mai 1983, l'Ecole est le seul Etablissement public universitaire professionnalisé du pays ayant pour mission de former des Techniciens Supérieurs, des Licenciés en informatique et des Ingénieurs informaticiens de haut niveau, aptes à répondre aux besoins et exigences d'informatisation des entreprises, des sociétés et des organismes implantés à Madagascar.

L'implantation de cette Ecole Supérieure de technologie de pointe dans un pays en voie de développement et dans une province à tissu économique et industriel faiblement développé, ne l'a pourtant pas empêché de former des spécialistes informaticiens de bon niveau recherchés par les sociétés et les organismes.

Depuis sa création jusqu'aujourd'hui, l'ENI placée sous la tutelle de l'Université de Fianarantsoa a formé et déversé sur le marché de l'emploi :

- 13 promotions d'Analystes Programmeurs, soit 447 diplômés.
- 22 promotions d'Ingénieurs Informaticiens, soit 554 diplômés.
- 13 promotions de Techniciens Supérieurs en Maintenance des Systèmes Informatiques, soit 310 diplômés.

Soit en tout 1311 diplômés.

La filière de formation d'Analystes Programmeurs a été créée en 1983, et a été gelée par la suite en 1996.

La filière de formation d'ingénieurs a été ouverte à l'Ecole en 1986.

La filière de formation de Techniciens Supérieurs en Maintenance des Systèmes Informatiques a été mise en place à l'Ecole en 1996 grâce à l'appui matériel et financier de la Mission Française de Coopération dans le cadre du Programme de Renforcement de l'Enseignement Supérieur (PRESUP).

Une formation pour l'obtention de la certification CCNA et/ou Network+, appelée « Cisco Networking Academy à Madagascar », a été créé en 2002-2003 grâce au partenariat avec Cisco System et l'Ecole Supérieure Polytechnique d'Antananarivo (ESPA). Cette formation n'existe plus actuellement.

Une formation doctorale a été ouverte à l'Ecole depuis l'année universitaire 2003-2004 grâce à la coopération entre l'Université de Fianarantsoa (ENI) et celle de Toulouse.

Enfin une formation en Licence Professionnelle en informatique ayant deux options : Administration des systèmes et des réseaux, Génie Logiciel et Base de Données, a été ouverte à l'Ecole pendant l'année universitaire 2007-2008.

La filière de formation de Techniciens Supérieurs en Maintenance des Systèmes Informatiques a été gelée en 2008.

Actuellement, l'administration de l'Ecole envisage de mettre en place les filières de formation de niveau MASTER.

1.4 Domaines de spécialisation

- Génie logiciel et Base de données,
- Administration des systèmes et des réseaux et
- Modélisation environnementale et Système d'Information Géographique.
- Intelligence artificielle et Traitement d'images

1.5 Architecture de la pédagogie

L'ENI forme des Licenciés et des Ingénieurs directement opérationnels au terme de leur formation respective. Ce qui oblige l'Ecole à entretenir des relations de collaboration étroites et permanentes avec les entreprises et le monde professionnel de l'Informatique à Madagascar.

La responsabilité de l'Ecole pour cette professionnalisation des formations dispensées implique de :

- Suivre les progrès technologiques et méthodologiques en Informatique (recherche appliquée, veille technologique, technologies Réseau, Multimédia, Internet,…).

- Prendre en considération dans les programmes de formation les besoins évolutifs des entreprises et des autres utilisateurs effectifs et potentiels, de la technologie informatique.

Cependant la professionnalisation ne peut se faire en vase close ; elle exige une «orientation client » et une « orientation marché ». Ce sont les entreprises qui connaissent mieux leurs besoins en personnel informatique qualifié. Ces entreprises partenaires collaborent avec l'ENI en présentant des pistes et des recommandations pour aménager et réactualiser périodiquement les programmes de formation. Ainsi, dans le cadre de ce partenariat avec les sociétés dans les divers bassins d'emploi en Informatique, l'Ecole offre sur le marché de l'emploi des cadres de bon niveau, directement opérationnels et avec des connaissances à jour.

L'architecture des programmes pédagogiques à l'Ecole s'appuie sur le couple théorie-pratique :

- Des enseignements théoriques et pratiques de haut niveau sont dispensés intra-muros à l'Ecole,
- Des voyages d'études sont effectués par les étudiants nouvellement inscrits et ayant passé une année d'études à l'Ecole,
- Des stages d'application et d'insertion professionnelle sont pratiqués en entreprise chaque année par les étudiants au terme de la formation académique à l'Ecole.

Les stages effectués en entreprise par les étudiants de l'ENI sont principalement des stages de pré-embauche.

Ces stages pratiques font assurer l'Ecole d'un taux moyen d'embauche de 97%, six mois après la sortie de chaque promotion de diplômés.

1.6 Filière de formation existantes et diplômes délivrés

Cycle Licence en informatique spécialisée en Administration des systèmes et des réseaux, en Génie logiciel et Base de données, aboutissant au Diplôme Universitaire de Licence. Les effectifs des étudiants au cours de l'année universitaire 2009 – 2010 sont les suivants :

- L1 (Première année de Licence) : 63,

- L2 (Deuxième année de Licence) : 53,
- L3 (Troisième année de Licence) : 52.

Cycle de formation d'Ingénieurs Informaticiens avec de compétences en Gestion, Systèmes et réseaux, de niveau Baccalauréat + 5 ans. Les effectifs des étudiants en 2009-2010 sont les suivants :

- ING2 (Deuxième année de la formation d'ingénieur) : 49,
- ING3 (Troisième année de la formation d'ingénieur) : 46.

La filière de formation en DEA en informatique organisée en partenariat avec l'Université Paul Sabatier de Toulouse, permet d'assurer une formation à la recherche par la recherche. Les trois meilleurs étudiants de la promotion effectuent généralement leurs travaux de recherche à Toulouse. Cette formation constitue un élément du système de formation de troisième cycle et d'études doctorales qui sera mis en place progressivement à l'ENI.

La thématique de recherche développée dans le Département de la Formation Doctorale porte sur la modélisation informatique et mathématique des systèmes complexes.

Une formation non diplômante en CISCO ACADEMY, soutenue par les Américains, avec certification CCNA existait à l'Ecole. Les effectifs des étudiants dans le système depuis sa création ont été les suivants :

- CISCO Première promotion 2002/2003 : 28,
- CISCO Deuxième promotion 2003/2004 : 5.

Le recrutement d'étudiants à l'ENI se fait chaque année uniquement par voie de concours d'envergure nationale, excepté pour le « Cisco Academy » et celui de la DEA, qui font l'objet de recrutement sur sélection des dossiers de candidature.

Bien qu'il n'existe pas au niveau international de reconnaissance écrite et formelle des diplômes délivrés par l'ENI, les diplômés de l'Ecole sont plutôt bien accueillis dans les Institutions universitaires étrangères. Des étudiants diplômés de l'Ecole poursuivent actuellement leurs études supérieures en 3ème cycle dans différentes Universités françaises et francophones, notamment à l'IREMIA de l'Université de La Réunion, à l'Université LAVAL au Canada, à l'Ecole Polytechnique Fédérale de Lausanne en SUISSE, à l'Ecole Doctorale STIC (Sciences et Technologies de l'Information et de la Communication) de l'Ecole Supérieure en Science Informatique de l'Université de Nice Sophia Antipolis.

1.7 Relations partenariales de l'ENI avec les entreprises et les organismes

1.7.1 Au niveau national

Les stages pratiqués chaque année par ses étudiants mettent l'Ecole en relation permanente avec plus de 300 entreprises, sociétés et organismes publics et privés nationaux et internationaux.

Parmi ces Etablissements, on peut citer : , ACCENTURE Maurice, AIR MAD, AMBRE ASSOCIATES, AUF, B2B, Banque Centrale, BFV SG, BIANCO, BLUELINE, BNI-CA, BOA, CEDII Fianar, CEM, Central Test, Centre Mandrosoa Ambositra, CNA, CNRIT, COLAS, COPEFRITO, Data Consulting, Pharmacie DES PLATEAUX Fianar, D.G. Douanes Tana, DLC, DTS/MOOV, FID, FTM, GNOSYS, IBONIA, IFIR des paramédicaux Fianar, INGENOSYA, INSTAT, IOGA, JIRAMA, Lazan'i Betsileo, MADADEV, MADARAIL, MAEP, MECI, MEF, MEN, MESupRES, MFB, MIC, MICROTEC, MININTER, MIN TéléCom et Nouvelles Technologies, NEOV MAD, NY HAVANA, OMNITEC, ORANGE , OTME, PRACCESS, QMM Fort-Dauphin, SECREN, SIMICRO, SNEDADRS Antsirabe, Société d'Exploitation du Port de Toamasina, Softewell, Strategy Consulting, TACTI, TELMA , ZAIN, WWF,…

L'organisation de stages en entreprise contribue non seulement à assurer une meilleure professionnalisation des formations dispensées, mais elle accroît également de façon exceptionnelle les opportunités d'embauche pour les diplômés.

Les diplômés de l'ENI sont recrutés non seulement par des entreprises et organismes nationaux, mais ils sont aussi embauchés dans des organismes de coopération internationale tels que l'USAID Madagascar, la Délégation de la Commission Européenne, la Banque Africaine de Développement (BAD), la Mission Résidente de la Banque Mondiale, la Commission de l'Océan Indien, etc.

1.7.2 Au niveau international

Entre 1996 et 1999, l'ENI a bénéficié de l'assistance technique et financière de la Mission Française de Coopération et d'Action Culturelle dans le cadre du PRESUP.

La composante du PRESUP consacrée à l'ENI a notamment porté sur :

- Une dotation en logiciels, microordinateurs, équipements de laboratoire de maintenance et de matériels didactiques ;
- La réactualisation des programmes de formation assortie du renouvellement du fonds de la bibliothèque ;
- L'appui à la formation des formateurs ;
- L'affectation à l'Ecole d'Assistants techniques français.

Et depuis le mois de mai 2000, l'ENI fait partie des membres de bureau de la Conférence Internationale des Institutions de formations d'Ingénieurs et Techniciens d'Expression Française (CITEF).

L'ENI a signé un Accord de coopération interuniversitaire avec l'IREMIA de l'Université de La Réunion, l'Université de RENNES 1 et l'Institut National Polytechnique de Grenoble (INPG).

Depuis le mois de juillet 2001, l'ENI abrite le Centre du Réseau Opérationnel (Network Operating Center) du point d'accès à internet de l'Ecole et de l'Université de Fianarantsoa. Grâce à ce projet américain financé par l'USAID Madagascar, l'ENI et l'Université de Fianarantsoa sont maintenant dotées d'une Ligne Spécialisée d'accès permanent à Internet. Par ailleurs, depuis 2002, une nouvelle branche de formation à vocation professionnelle a pu y être mise en place, en partenariat avec Cisco System.

Enfin et non de moindres, l'ENI a noué des relations de coopération avec l'Institut de Recherche pour le Développement (IRD). L'objet de la coopération porte sur la Modélisation environnementale du corridor forestier de Fianarantsoa. Dans le même cadre, un atelier scientifique international sur la modélisation des paysages a été organisé à l'ENI au mois de Septembre 2008.

Comme l'ENI constitue une pépinière incubatrice de technologie de pointe, d'emplois et d'entreprises, elle peut servir d'instrument efficace pour la lutte contre la pauvreté.

De même que l'Ecole permet de renforcer la position concurrentielle de la Grande île sur l'orbite de la mondialisation grâce au développement des nouvelles technologies.

1.8 Ressources humaines

- Directeur : Monsieur Fontaine RAFAMANTANANTSOA,
- Chef du Département de la Formation Théorique : Monsieur Venot RATIARSON,

- Chef du Département de la Formation Pratique : Monsieur Cyprien RAKOTOASIMBAHOAKA,
- Chef du Département de la Formation Doctorale : Monsieur Josvah Paul RAZAFIMANDIMBY,
- Nombre d'Enseignants permanents : 12,
- Nombre d'Enseignants vacataires : 10,
- Personnel administratif et technique : 15.

1.9 Projets et perspectives de développement institutionnel (2010-2014)

- Restructuration du système pédagogique de l'Ecole selon le schéma LMD (Licence-Master-Doctorat),
- Mise en place à l'Ecole d'un Département de Formation de 3ème cycle et d'études doctorales en Informatique,
- Création à l'Ecole d'une Unité de Production Multimédia, d'un club de logiciel libre Unix,
- Projet TICEVAL, « Développement et certification de compétences TIC : transfert et adaptation d'un dispositif humain et technique », projet mené par l'Université de Savoie en collaboration avec l'ENI de Fianarantsoa sur financement du Fonds Francophones des Inforoutes. Période d'exécution du projet : 2010-2012.

Chapitre 2

Présentation de la société BLUELINE

Blueline est un fournisseur d'accès Internet, elle est le partenaire idéal pour les réseaux de communications professionnelles. Aux particuliers, Blueline offre le meilleur de l'Internet.

Sa vocation est de fournir des solutions de transmission de données pour les réseaux d'entreprises, l'Internet à haut débit, la connexion par modem avec assistance complète pour les particuliers.

Par ailleurs, Blueline innove dans la façon de vulgariser Internet en proposant le service mail gratuit et le surf gratuit dans ses locaux. En cela, Blueline s'attache fortement aussi à son rôle moteur dans la diffusion des NTIC.

2.1 Société

Créée en 1997, la société Blueline est le pionnier de l'accès à l'Internet et des transmissions de données pour les entreprises à Madagascar.

Blueline propose à ses clients une expérience de plus de 10 ans dans les réseaux télécom à Madagascar et conçoit des offres sur mesure pour ses clients de grands comptes.

Ainsi, Blueline propose la meilleure des technologies basées sur des solutions de transmission de données :

- Terrestre :
  - WIMAX,
  - Faisceau Hertzien,
  - Laser,
  - Fibre Optique.
- Satellitaires:
  - VSAT Band Ku,
  - VSAT Band C.

En s'appuyant sur le savoir faire de Gulfsat Madagascar en termes d'ingénierie télécom et connaissance des réseaux IP (maison mère de Blueline), Blueline développe des offres adaptées à l'ensemble des entreprises :

- Internet,
- Liaisons point à point nationales et internationales,
- Réseaux privé.

2.2 Ressources humaines

- Directeur général : Damien DE LAMBERTERIE
- Directeur des opérations : Andriamampandry RAZANAKOTO
- Directeur Technique : Jérôme SANTINI
- Directeur Administratif et Financier : Ndrianja RAJEMISON

2.3 Adresse

Société Blueline
Route Digue
Domaine d'Andranoabo
BP : 1484
101 Tananarive MADAGASCAR

2.4 Organigramme

La **Figure 2: Organigramme simplifié de la société Blueline**, *cf. page 17*, présente l'organigramme simplifié de la société.

2.5 Produits de Blueline

La société Blueline propose plusieurs produits tels que :

- Hébergement Web

Blueline propose des solutions d'hébergement adaptées aux besoins des clients en terme de sécurité, continuité, rapidité, fiabilité.

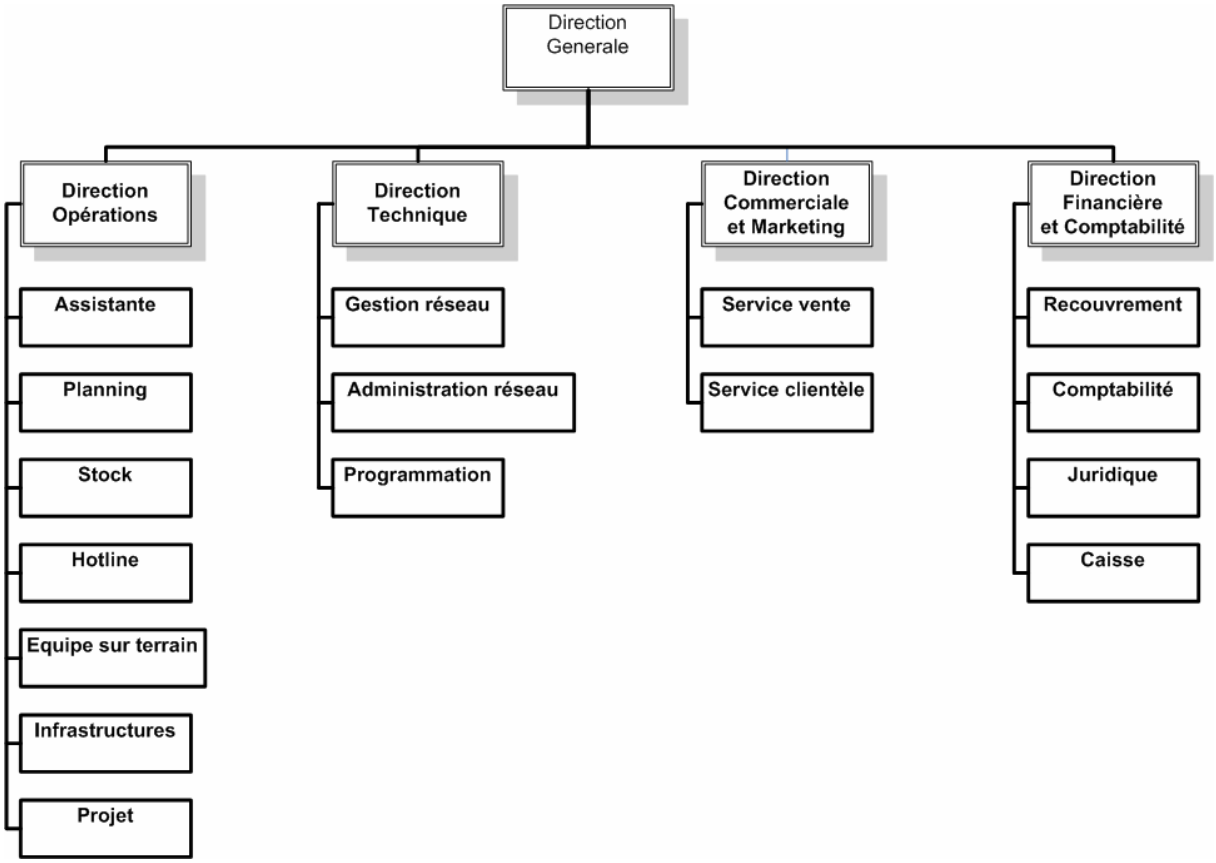

Figure 2: Organigramme simplifié de la société Blueline

- Nom de domaine

Ainsi, Blueline peut donner comme vitrine sur Internet un site accessible depuis :

- http://www.votre.societe.mg pour Madagascar,
- http://www.votre.societe.com ou http://www.votre societe.net pour l'étranger.

Et de plus, dans le contexte de forte évolution que connaît le secteur télécom à Madagascar, Blueline innove en élargissant sa gamme de produits Grand Public et annonce la première offre *Triple Play* à Madagascar :

- Une offre internet basée sur la toute dernière génération de réseau 4G,
- Une offre de téléphonie mobile,
- Une offre de télévision incluant les plus grandes chaînes Malgaches et Internationales,
- Une offre Triple Play associant Internet, téléphonie mobile et télévision dans un contrat unique.

Blueline réaffirme sa volonté de jouer un rôle majeur à Madagascar pour permettre au plus grand nombre de disposer du meilleur des technologies modernes : Internet, Téléphonie Mobile et TV.

2.5.1 Triple Play: 4G, Phone, TV

A l'égal des pays les plus avancés en termes d'intégration High-Tech destinée au grand public, Blueline est le premier opérateur de Madagascar à proposer une offre *Triple Play incomparable* : Internet 4G, Téléphonie Mobile et TV *à partir de 79 000 Ar TTC* par mois*.

Cette offre permet d'optimiser le budget du foyer en répondant aux attentes de chacun des membres de la famille.

4G

A travers les investissements réalisés par Blueline ces derniers mois, Madagascar est un des tous premiers pays au monde à disposer, dès cette année, d'une offre commerciale Internet basée sur un nouveau réseau de 4ème génération.

Il est incontestable que le véritable accès au Très Haut Débit n'est possible qu'au travers de technologies d'accès adaptées (réseau 100% IP) : en la matière, la 4G fait l'unanimité de tous les experts du secteur.

Phone

Blueline Phone est une offre de téléphonie mobile *nationale* tout particulièrement optimisée pour permettre d'appeler ses correspondants sans compter, tout en maîtrisant son budget. En effet, tous les bénéficiaires de l'offre Blueline Phone Access peuvent s'appeler et s'envoyer des SMS entre eux en *illimité.*

TV

Blueline TV est conçue pour permettre au plus grand nombre d'accéder au meilleur des chaînes Malgaches et Internationales.

Le 1er Pack Blueline TV Access est proposé sans facturation récurrente pour faciliter l'ouverture aux contenus internationaux.

Quatre autres Packs donneront l'opportunité de visualiser des dizaines de chaînes dans différentes langues, traitant d'actualités, de culture, de sport et bien d'autres sujets… L'offre sera disponible courant septembre 2010.

2.6 Equipe

Blueline maîtrise l'espace avec deux locaux sis à Ankorondrano et à la Digue. Celui d'Ankorondrano regroupe les ressources administratives et commerciales tandis que celui de la Digue centralise les services techniques et clientèles.

Plus de 80 ingénieurs et techniciens supérieurs spécialisés avec des compétences diverses aussi bien en télécommunication qu'en informatique associée à des jeunes cadres maîtrisant les aspects organisationnels et marketing, Blueline est très réactive face au marché et aux mutations technologiques.

2.6.1 Compétences

Blueline déploie ses atouts dans des domaines parfaitement connus.

On peut citer:

- Plateau technique :
  - Hébergement et gestion de sites Web à forte valeur ajoutée.
- Topologie et réseau de communication :
  - Interconnexion de sites distants, connexions partagées, liaisons spécialisées, routage de communication, travail collaboratif, visioconférence,
  - Installation de réseaux,
  - Fourniture d'accès Internet sous toutes ses formes.

2.7 Vision

Blueline se positionne comme l'opérateur disposant des meilleurs des infrastructures disponibles à Madagascar. Selon le niveau d'exigence et les enjeux exprimés par les clients, la société Blueline cherche à optimiser les solutions Télécom à disposition.

Ainsi, par une écoute attentive, Blueline accorde une importance toute particulière pour comprendre les projets de ses clients, les accompagner dans l'expression des besoins si nécessaire, et apporter les solutions qui répondent à des enjeux.

Ces enjeux peuvent prendre en compte les aspects suivants :

- L'optimisation des coûts
  - Suivant le choix d'architecture,
  - Suivant un mix adapté des CAPEX/OPEX de la solution.
- La qualité de service :
  - Stabilité des débits,
  - Disponibilité plus ou moins critique par site,
  - Contrôle des services mis en œuvre,
  - Engagements (SLA) tenant compte des contraintes.
- L'évolutivité de la solution
  - Prise en compte des évolutions technologiques,
  - Adaptation de l'architecture en fonction des volumes actuels et à venir,
  - Capacité à délivrer un service quelque soit la contrainte géographique.
- La flexibilité de la solution
  - Prise en compte des évolutions de périmètre en cours de contrat.
- Une durée contractuelle adaptée à des projets.

Pour ce faire, la société Blueline s'efforce de concentrer l'ensemble de ses services sur une gestion optimale des besoins Télécom :

- pour permettre aux clients de se concentrer sur le coeur de métier,
- les accompagner lors de choix stratégiques Télécom,
- les considérer comme un partenaire de confiance.

La confiance guide toutes les activités de la société.

2.8 Gestion du projet

Une des clés de la satisfaction de clients réside dans la capacité de la société à gérer des projets complexes.

Ainsi, les moyens mis en oeuvre pour la réussite des projets se basent sur l'expérience de collaborateurs et sur une méthodologie dont les principales caractéristiques sont les suivants:

- Mise en place d'une équipe technico-commerciale composée de managers expérimentés,
- Définition des enjeux pour les clients,
- Evaluation des choix techniques possibles,

- Proposition répondant à des enjeux,
- Déploiement de la solution :
    - Désignation de l'équipe responsable du projet,
    - Comités de pilotage réguliers,
    - Mise en commun des informations relatives au déploiement,
    - Respect des délais,
    - Tests et Recette de la solution.

- Gestion opérationnelle de la solution,
- Mise en place de la structure de support client (Customer Care),
- Mise en place de l'équipe de suivi des opérations,
- Suivi commercial permanent.

2.9 Support client

Les équipes de support client sont organisées en fonction de la complexité du problème technique à résoudre :

- Une hotline téléphonique (24h/24,7j/7) permettant de résoudre immédiatement 95% des demandes clients,
- Une équipe d'analystes réseaux permettant d'appréhender les cas les plus complexes,
- Une équipe de techniciens terrain pour résoudre physiquement les éventuelles pannes des infrastructures réseaux,
- Un point de contact unique identifié pour assurer un meilleur suivi des besoins des clients.

2.10 Réseaux

Blueline maîtrise l'ensemble de ses infrastructures au travers de ses équipes techniques et opérationnelles internalisées.

Elle dispose de réseaux de collecte terrestres et satellitaires nationaux interconnectés à deux câbles en fibre optique Lion et Eassy.

Blueline est à ce jour le seul opérateur à disposer de cette double connectivité garantissant un haut niveau de performance et de disponibilité.

2.11 Supervision

Blueline effectue un contrôle de l'ensemble de ses infrastructures depuis le coeur du réseau jusqu'aux équipements installés chez le client, ce qui garantit :

- Un taux de disponibilité très élevé des systèmes, proche de 100% pour l'ensemble des liaisons,
- Une mise en oeuvre d'engagements de services (SLA) adaptés à la demande du client,
- Une optimisation des interventions terrain grâce à une équipe technique présente dans les principales villes de province.

2.12 Carte du Backbone BLUELINE

La figure ci-dessous montre la carte du Backbone BLUELINE. Les différents sites sont reliés soit par une liaison en fibre optique, soit par un faisceau hertzien.

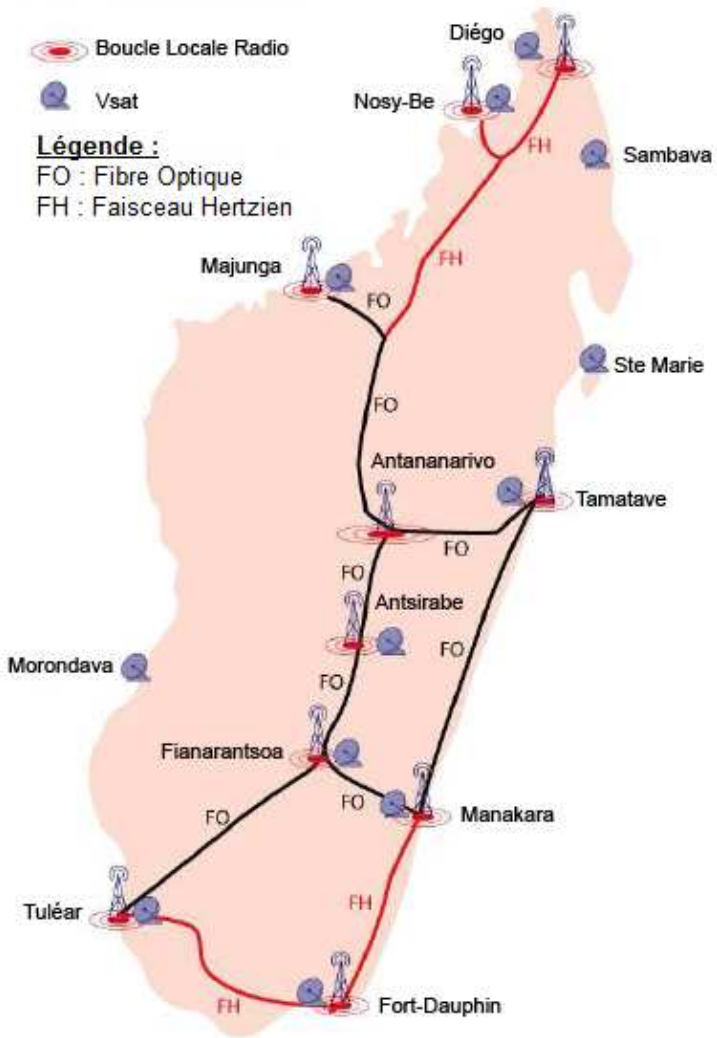

Figure 3: Carte du Backbone Blueline

Partie II

Analyse préalable

CHAPITRE 3 :

Description du projet

CHAPITRE 4 :

Analyse de l'existant

Chapitre 3

Description du projet

Avant de rentrer dans le vif du sujet, ce chapitre fournira une description brève et précise du sujet.

3.1 Contexte

Au sein d'un réseau informatique, l'activité des serveurs évolue en fonction :

- des besoins,
- des droits d'accès,
- des accès de plus en plus restrictifs,
- des accès de plus en plus rapides.

Ces facteurs influent non seulement sur les ressources système, mais surtout sur l'espace disque des serveurs. En effet, la masse de données à gérer augmente constamment, le besoin en capacité de stockage augmente en conséquence.

L'infrastructure classique de stockage arrive à ses limites. Si un serveur atteint sa limite en termes de performances, il peut toutefois, et sous certaines conditions, continuer à fonctionner avec des performances réduites et il sera possible de repousser l'investissement de quelques mois. Mais, si le système de stockage atteint sa capacité maximale, il faudra investir. Cela signifie aussi que les efforts mis en place pour gérer les données vont continuer à augmenter.

Par ailleurs, les moyens financiers restent limités pour satisfaire ce grand besoin en capacité.

Il faut donc trouver des solutions permettant de répondre aux nouvelles demandes de gestion de la masse de données tout en accélérant les accès à ces différentes données.

3.2 Problématique

Cette situation est aujourd'hui une réalité au sein des entreprises, y compris la société BLUELINE, les administrateurs systèmes, et les responsables des systèmes d'informations sont face à plusieurs interrogations :

- Comment répondre aux besoins de stockage face à la croissance de la masse de données, sans surdimensionner le réseau ?
- Comment assurer la conservation et l'intégrité des données ?
- Comment répondre aux besoins croissants d'un accès plus rapide à ces données ?

Ainsi, avant d'investir, le responsable doit mener une profonde analyse des possibilités offertes, afin de choisir la technologie adaptée.

Ce mémoire s'appuiera sur l'hypothèse selon laquelle les réseaux de stockage SAN (Storage Area Network) et/ou NAS (Network Attached Storage) peuvent aider à gérer plus facilement et de façon plus économique cette masse croissante de données.

Coïncidant à notre demande de stage, la société BLUELINE nous a proposé d'étudier les différentes technologies pour résoudre les problèmes suscités.

Il faut donc trouver la meilleure solution, en tenant compte de la complexité de mise en place et d'administration.

3.3 Méthodologie d'étude

Pour ce faire, le chapitre suivant présente l'analyse de l'existant au sein de la société BLUELINE afin de dégager ses contraintes et ses limites.

Vient ensuite, l'étude approfondie des technologies à la résolution des questions liées au partage et au stockage des données.

Et la solution proposée fera par la suite l'objet :

- d'un point de vue conceptuel, suivi
- d'une réalisation technique.

Chapitre 4

Analyse de l'existant

Une bonne compréhension de l'architecture réseau existante aide à déterminer la portée du projet et l'implémentation de la solution. Il est essentiel de disposer d'informations précises sur l'infrastructure réseau et les problèmes qui ont une incidence sur le fonctionnement du réseau. En effet, ces informations affectent une grande partie des décisions que nous allons prendre dans le choix de la solution et de son déploiement.

4.1 Architecture réseau globale

La configuration globale de l'architecture réseau de la société Blueline est présentée sur la **Figure 4: Architecture réseau globale**, *cf. page 27*.

Les deux câbles Eassy et Lion relient la société vers l'Internet. Ces deux câbles sont par la suite reliés à un dispatcher avant d'arriver sur le routeur principal.

Les serveurs « proxy squid », le serveur « cache P2P » et le « backbone » interne sont connectés directement à ce dernier. Sur le « backbone » interne est connecté :

- Un switch pour les serveurs et
- Les liaisons Faisceaux Hertziennes vers les différents points d'accès.

Remarque :

Il n'existe qu'un seul réseau pour tous les serveurs.

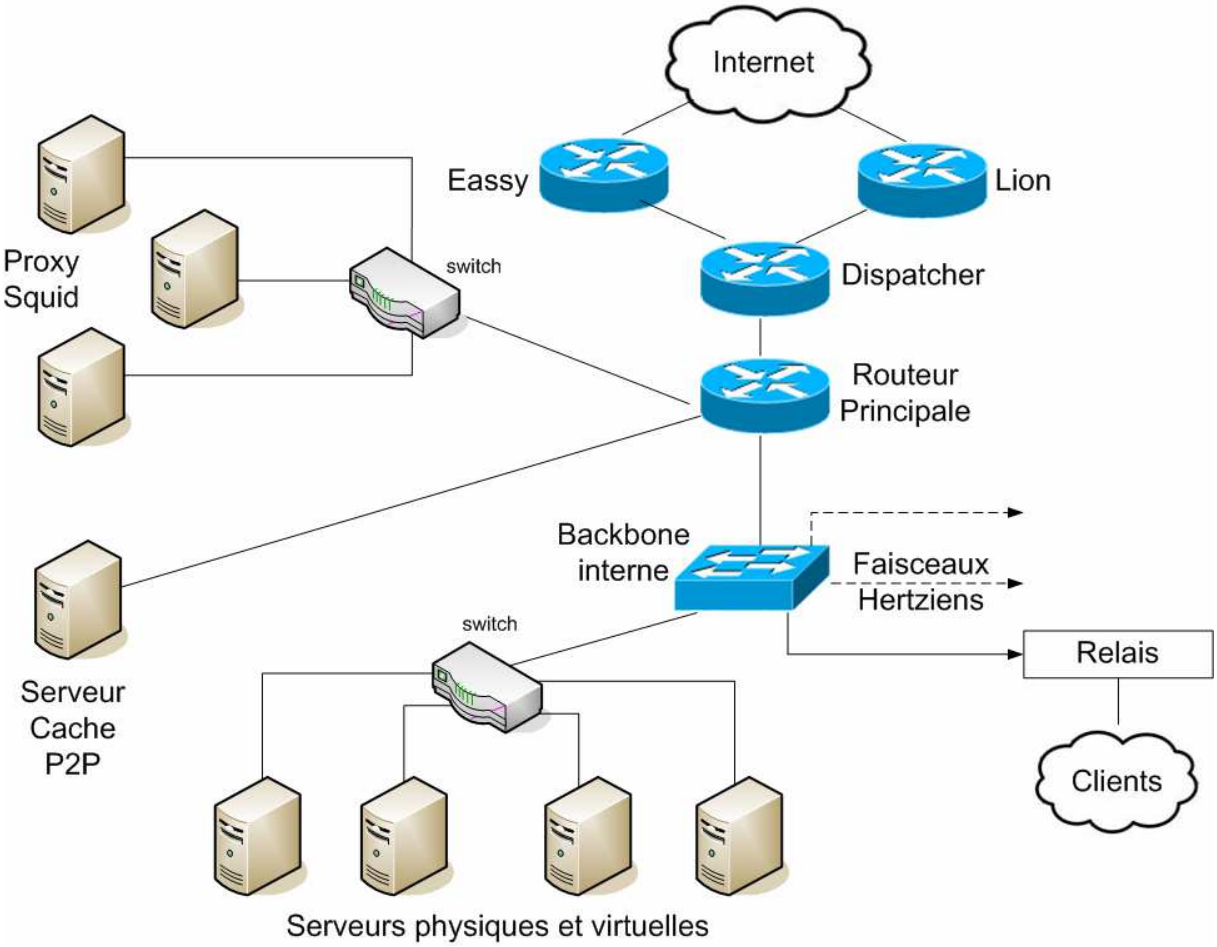

Figure 4: Architecture réseau globale

4.2 Architecture réseau des serveurs

Puisque le projet se focalise sur les serveurs, la **Figure 5: Architecture réseau des serveurs**, *cf. page 28*, montre l'architecture réseau simplifiée des serveurs.

Cette figure montre que :

- le stockage des données est encore local sauf pour le mail qui utilise un système de stockage réseau avec NFS,
- le serveur NFS est un SPOF (Single Point Of Failure), c'est-à-dire que s'il tombe en panne, il n'a pas d'autres serveurs pour prendre le relais.

Les serveurs utilisés sont soit des serveurs physiques, soit des serveurs virtuelles en openvz ou en KVM.

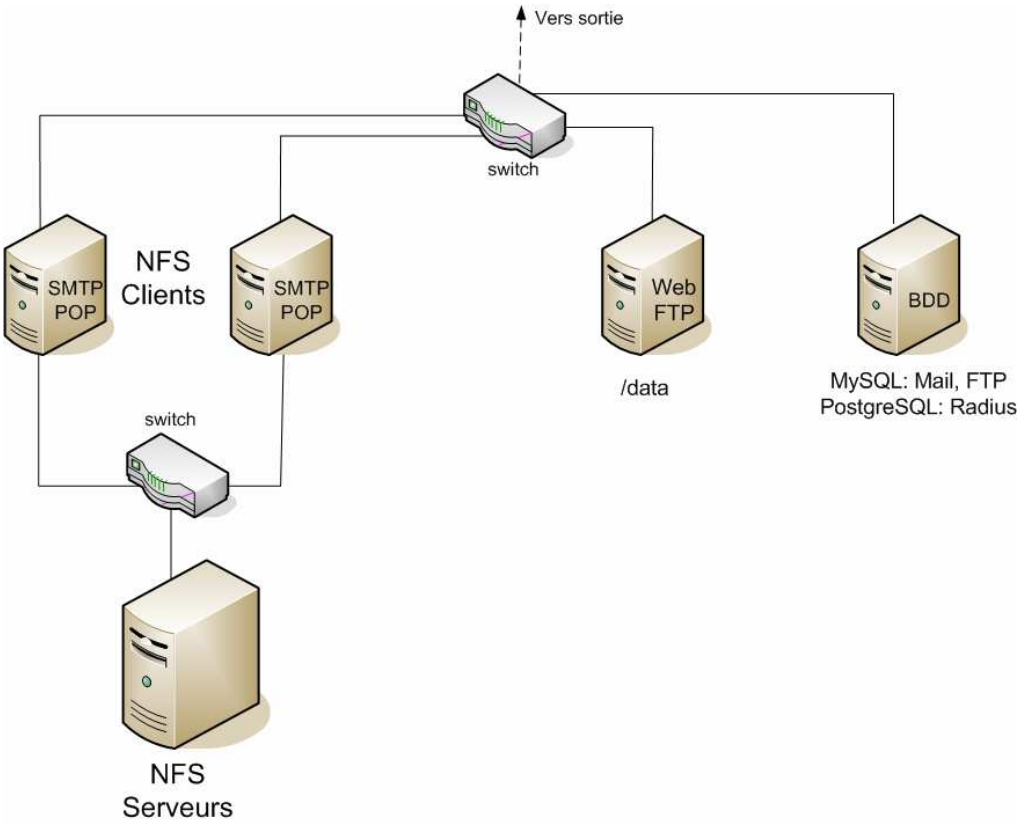

Figure 5: Architecture réseau des serveurs

Les serveurs virtuels sous conteneurs openvz sont:

- ciscolog : serveur de log,
- dbmail : BDD des comptes mails,
- NS serveurs,
- serveur TFTP,
- serveur VPN pour la connexion entre les sites clients,
- serveur RADIUS pour l'authentification,
- divers serveurs SMTP pour l'envoi de mail en sortie,
- MX secondaire,
- pg-a : base de données POSTGRES pour certain applications et
- serveurs de monitoring interne.

Les serveurs virtuels en KVM sont :

- MAS1 et MAS2 : serveur d'authentification aptilo,
- DHCP pour 4G et
- selfcare : serveur portail pour client 4G.

Les serveurs physiques sont :

- mailaka data: serveur de données pour mail,
- serveurs SMTP relais pour les clients et serveur POP/IMAP des clients et
- serveur WEB et FTP.

4.3 Avantages et inconvénients

Les serveurs virtuels ont pour avantage :

- d'optimiser les ressources des machines,
- de réduire la consommation en électricité,
- de faciliter les maintenances des machines : cas du mise à jour du noyau,
- de faciliter son déploiement : il suffit de copier le snapshot d'une machine pour en créer un autre,
- de réduire le temps de panne puisqu'il est possible de dupliquer facilement une machine virtuel vers un autre serveur.

D'autre part, cette virtualisation présente des inconvénients :

- le système de stockage par NFS ne fonctionne pas avec openvz,
- il n'est pas encore testé avec KVM, même si cela marche, on ne bénéficie pas du partage de ressource avec openvz.

4.4 But du projet

Vu les avantages et les inconvénients suscités de l'architecture actuelle, ce projet a pour but de mettre en place un système de stockage distribué et si possible pouvant fonctionner avec le système openvz.

Ainsi on pourrait virtualiser seulement les applications web, ftp, mail, base de données, etc. et garder les données : boites mail, répertoire web et ftp, données de la base de données sur un serveur de stockage.

On pourrait facilement améliorer le système de stockage par rapport à l'extension, la fiabilité et la disponibilité.

Partie III

Analyse conceptuelle

CHAPITRE 5 :

Réseaux SAN/NAS

CHAPITRE 6 :

Architecture réseau de la solution retenue

Chapitre 5

Réseaux SAN/NAS

Cette partie présente les différentes solutions proposées en matière de stockage en réseau, les réseaux SAN/NAS, afin de connaître celle qui réponds aux attentes de la société BLUELINE.

5.1 Architecture DAS (Direct Attached Storage)

5.1.1 Définition

Le DAS (Direct Attached Storage) est l'architecture standard de stockage avec connexion directe aux serveurs. C'est encore l'architecture la plus courante aujourd'hui. L'hôte et ses éléments de stockage sont connectés un à un par des liens SCSI.

5.1.2 Principe de fonctionnement

Le stockage DAS est un dispositif interne ou externe qui se branche directement à un seul serveur. Ce dispositif contient des disques indépendants de type RAID tolérant la panne de disques sans perte de données et qui fonctionne de façon indépendante des systèmes d'exploitation.

Dans cette topologie de stockage, à un serveur correspond un équipement de stockage. Aucune fonctionnalité du réseau n'est alors déployée.

Par conséquent, pour accéder à ce dispositif, les usagers doivent accéder au serveur auquel il est branché.

5.1.3 Architecture

Le DAS est l'architecture standard de stockage avec connexion directe aux serveurs. Les postes de travail individuels trouvent grâce à la connexion sur le réseau, la possibilité de mettre en place l'échange de données et le travail coopératif.

Un tel réseau est représenté par la figure suivante :

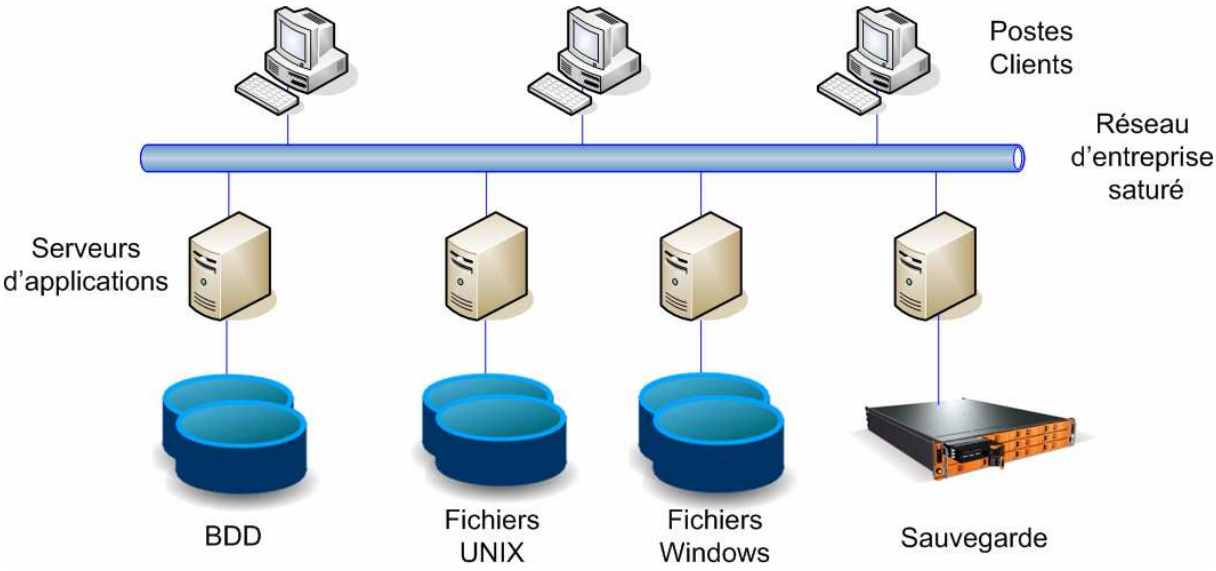

Figure 6: Architecture DAS

Cette figure montre qu'il est difficile d'optimiser le stockage.

Le réseau s'est étendu, mais les principes qui régissent les échanges de données restent identiques. Les données sont stockées sur disque dur, l'accès à ces données est piloté par le serveur local. Le partage ou l'échange de données entre les machines, est basé sur un protocole client/serveur (NFS ou FTP). Sollicité par un client, le serveur est le passage obligé de tous les flux concernant ses espaces disques locaux.

L'évolution naturelle de cette architecture est le serveur de fichiers. Elle consiste à dédier un gros serveur aux fonctions de stockage. Les aspects spécifiques sont développés sur ce serveur pour une meilleure efficacité et un meilleur service.

5.1.4 Avantages et inconvénients

Avantages

Il est très simple de mettre en place un DAS puisqu'il n'y a qu'un serveur branché au dispositif de stockage.

Chaque serveur réserve la totalité de ses ressources aux applicatifs qu'il héberge. La fonction de gestion des données, centrée sur un tel serveur, permet une économie de moyens (matériels et humains) et un service plus souple et plus efficace.

De plus, dans cette configuration, on tire aussi les avantages classiques d'une solution centralisée:

- Une réponse efficace à l'augmentation des besoins en espace disque,
- Une mise en place plus avisée de la sécurité,
- Une intégration de matériels fiables (des éléments matériels redondants, la technologie RAID, …),
- Un partage des données plus aisé.

Cependant, du choix de la configuration dépende directement les performances d'entrées/sorties, les possibilités d'extension et la sécurité de données. Des moyens matériels appropriés doivent être intégrés afin d'assurer la meilleure disponibilité des données. Le serveur devient donc un point central de l'architecture, et est indispensable à la bonne marche des applications.

Inconvénients

Le débit du réseau devient rapidement un point sensible de l'architecture, mais ce n'est pas le seul. Le protocole SCSI limite le nombre de périphériques connectés, la performance de la machine hôte et les applications hébergées ont un impact direct sur les performances des entrées/sorties. Une panne de la machine, entraîne l'indisponibilité des données.

Cette topologie rend difficile les évolutions (en volume ou en débit). Cette architecture a tendance à être limitée aux configurations de taille réduite.

5.2 Solution NAS

5.2.1 Définition

Un NAS (*Network Attached Storage*) est un dispositif de stockage en réseau. Il s'agit d'un serveur de stockage à part entière pouvant être facilement attaché au réseau de l'entreprise afin de servir de serveur de fichiers et fournir un espace de stockage tolérant aux pannes.

5.2.2 Principe de fonctionnement

Un NAS définit un produit spécifique possédant sa propre adresse IP (statique) et directement connecté au réseau local de l'entreprise. Par conséquent, l'installation d'un tel

matériel ne nécessite pas la mise en place d'une infrastructure spécifique. C'est un avantage en terme de coût, de temps et de main d'œuvre surtout pour les PME.

Le NAS se distingue par les caractéristiques suivantes :

- Directement connecté au réseau (il n'est pas nécessaire de passer par le serveur d'application pour ajouter du stockage) et permet l'accès client aux données sans passer par un serveur d'application.
- Supporte le partage de fichiers dans des réseaux hétérogènes sous Windows, Apple, ou les systèmes d'exploitation basés sur Unix.
- Utilise un système d'exploitation et un système de fichiers dédiés qui résident dans le serveur NAS et qui sont gérés par ce dernier.
- Relativement facile à déployer et à gérer.

L'utilisation d'un NAS est adaptée aux applications faisant appel au service de fichiers comme l'hébergement de sites WEB ou encore les serveurs de fichiers ou de messagerie.

5.2.3 Architecture

Les serveurs de fichiers tendent à disparaître au profit des NAS, plus simples à sauvegarder, plus adaptables aux besoins et plus facilement administrable. De plus, un seul serveur de stockage est maintenant nécessaire puisque les NAS supportent plusieurs types de systèmes de fichiers (*cf.* **Figure 7: Architecture NAS**, *page 35*).

5.2.4 Avantages et inconvénients

Avantages

L'avantage, de cette solution, est qu'une unité de stockage est autonome. La technologie de stockage sur réseau (NAS) est reconnue pour la facilité de sa gestion et sa possibilité d'utiliser des réseaux Ethernet, à moindre coût, pour partager les fichiers.

L'installation est relativement rapide et la capacité de stockage est attribuée aux usagers sous une base de demande.

Elle est évolutive grâce à des possibilités de branchement de nouveaux disques à chaud. La redondance des données se fait au sein d'une seule unité de stockage, grâce à la technologie RAID.

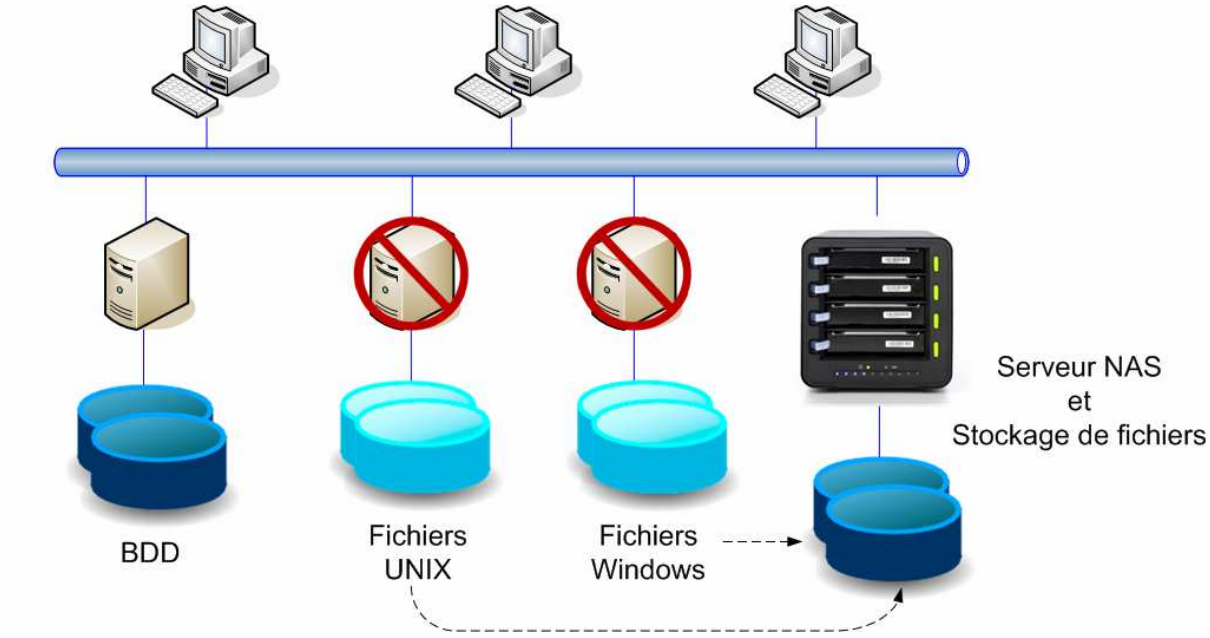

Figure 7: Architecture NAS

Les serveurs de fichiers tendent à disparaître au profit des NAS, plus simples à sauvegarder, plus adaptables aux besoins et plus facilement administrable. De plus, un seul serveur de stockage est maintenant nécessaire puisque les NAS supportent plusieurs types de système de fichiers.

Inconvénients

L'inconvénient principal, de ce type de solutions, est qu'elle vient surcharger le réseau de l'entreprise comme le montre le schéma ci-dessus.

Ensuite, le NAS s'oriente plus dans le cadre d'un serveur de fichiers autonome. Et le fait de passer par mode de communication non dédié comme TCP/IP ralentit le débit et le temps de réponses d'accès aux données. Pour des serveurs d'applications, avec des traitements lourds et conséquents, ce n'est peut être pas la véritable solution.

De plus, le NAS utilise un transfert en mode fichier, les performances de transport de gros fichiers sont rapidement limitées.

5.3 Solution SAN

5.3.1 Définition

Un SAN (Storage Area Network) est un réseau de stockage à part entière.

Le SAN est basé sur un réseau très haut débit en Fibre Channel ou SCSI, des équipements d'interconnexion dédiés (switch, ponts) et des éléments de stockage (disques durs) en réseau.

5.3.2 Principe de fonctionnement

Basé sur la Fibre Channel, la topologie indépendante et multicouche fonctionnant en série et se comportant exactement comme une liaison téléphonique, le SAN est un réseau de stockage ouvert et évolutif qui relie des périphériques de stockage, des serveurs/stations et des postes de travail, par ailleurs reliés au réseau d'entreprise.

Le SAN constitue une plate-forme de communication qui exploite les protocoles SCSI sur des technologies d'interconnexion à haut débit. Il virtualise totalement l'espace de stockage et travaille au niveau des blocs (et non des fichiers comme les serveurs NAS); ceci permet le partage centralisé des données via des switchs intelligents Fibre Channel.

Un SAN se différencie des autres systèmes de stockage tel que le NAS par un accès bas niveau aux disques. Pour simplifier, le trafic sur un SAN est très similaire au principe utilisé pour l'utilisation des disques internes (ATA, SCSI). C'est une mutualisation des ressources de stockage.

Dans le cas du SAN, les baies de stockage n'apparaissent pas comme des volumes partagés sur le réseau. Elles sont directement accessibles en mode bloc par le système de fichiers des serveurs. En clair, chaque serveur voit l'espace disque d'une baie SAN auquel il a accès comme son propre disque dur. L'administrateur doit donc définir très précisément les LUN (unités logiques) et le zoning, pour qu'un serveur Unix n'accède pas aux mêmes ressources qu'un serveur Windows utilisant un système de fichiers différent.

Un SAN se distingue par les caractéristiques suivantes :

- Permet aux serveurs un accès partagé à une ferme de stockage commune et à une ou plusieurs librairies de bandes pour la sauvegarde et la restauration,
- Utilise un réseau Fibre Channel séparé spécifique au stockage,
- Assure les transferts de données stockées entre les serveurs et les dispositifs de stockage sur le SAN, allégeant de ce fait la charge du LAN,
- Permet l'installation distante de sous-systèmes de disques durs et de librairies de bandes.

5.3.3 Architecture

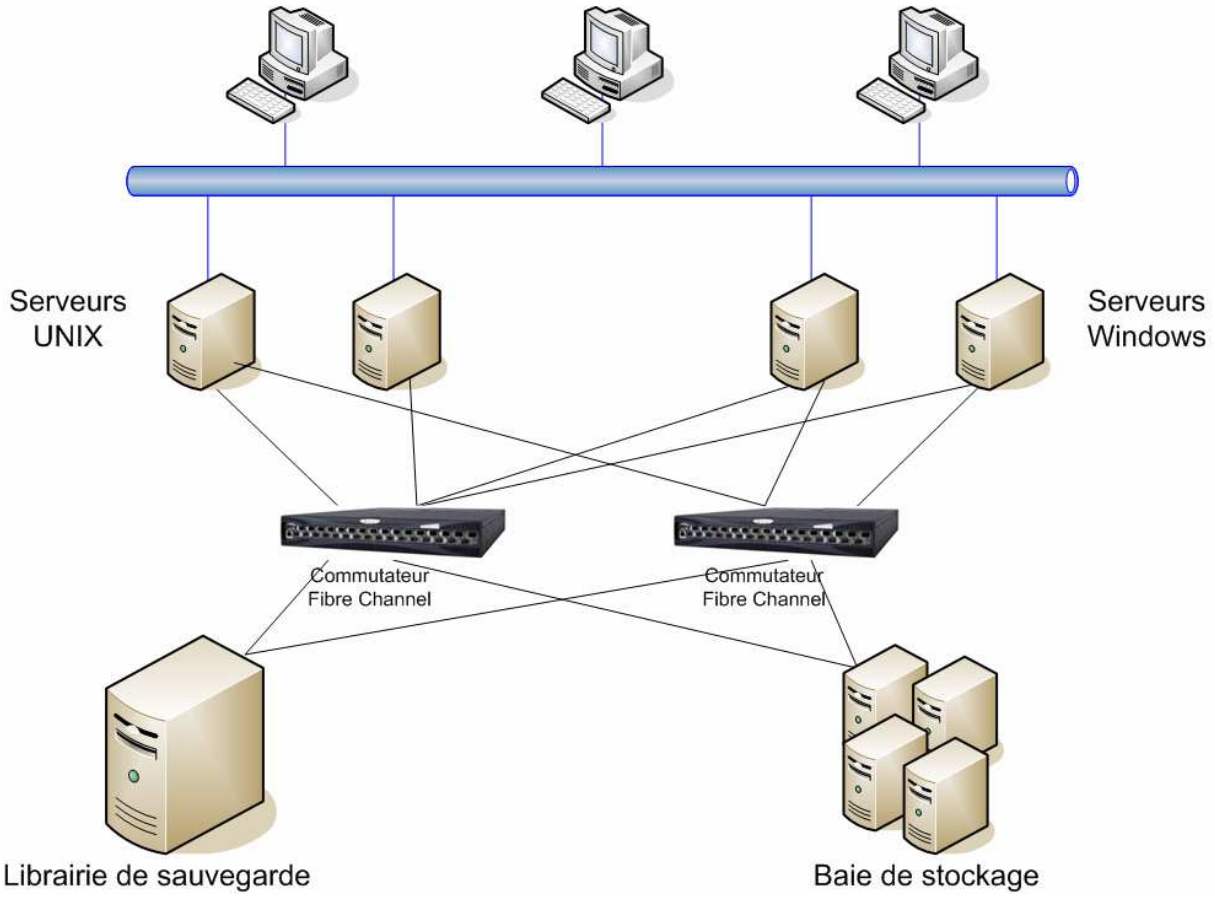

Figure 8: Architecture SAN

5.3.4 Constitution

Le réseau SAN est basé sur le protocole Fibre Channel (FC). Le SAN est un réseau sur lequel sont connectés des serveurs et des périphériques de stockage. Chaque serveur peut accéder à chaque périphérique.

La majorité des SAN sont basés sur un réseau de transport en fibre optique, mais on peut aussi utiliser le cuivre. Le protocole utilisé pour le transfert des données est le SCSI série (SCSI-3).

L'implantation actuelle autorise des débits de 100Mo/s, des distances jusqu'à 10 Km, et des centaines de périphériques sur un même réseau.

Trois topologies ont été définies :

- La topologie « point à point » : C'est la topologie la plus simple qui offre la meilleure bande passante.

Figure 9: Topologie point à point

Les configurations « point à point » sont souvent les plus anciennes. Le protocole était alors limité à 25Mo/s à cause des performances des serveurs et des disques. Les configurations « point à point » sont encore tout à fait adaptées à des environnements simples, bien que l'augmentation des débits, et le développement des switchs FC aient favorisé l'émergence des deux autres topologies.

- La topologie « en anneau » peut connectée jusqu'à 126 éléments entre eux. Elle reprend les principes du protocole Token Ring, ou l'accès au réseau est arbitré entre chaque périphérique du réseau. La bande passante de la boucle est partagée entre chaque périphérique du réseau.

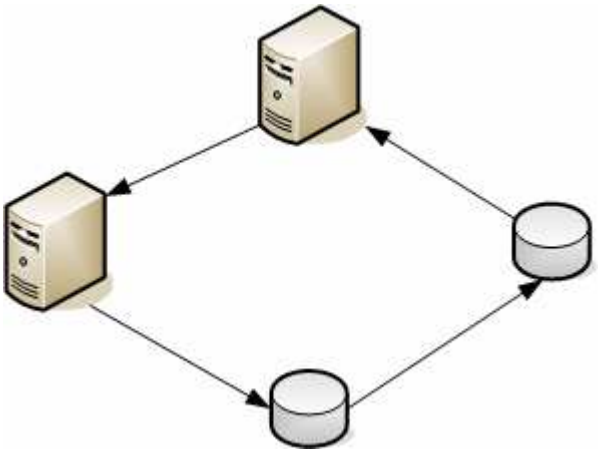

Figure 10: Topologie en boucle

Les premières boucles ont été réalisées comme l'indique la figure ci-dessus. Si un composant est défaillant, la boucle est cassée et les composants de la boucle sont hors circuit.

Les boucles sont actuellement réalisées à l'aide d'un hub qui réalise une topologie en étoile pour une configuration en boucle. Ceci permet de mettre hors circuit un élément défaillant sans casser la boucle.

Cette topologie est assez souple. Elle permet la connexion de 126 périphériques. Elle est relativement moins coûteuse, mais impose le partage de bande passante.

Elle est représentée par la figure suivante :

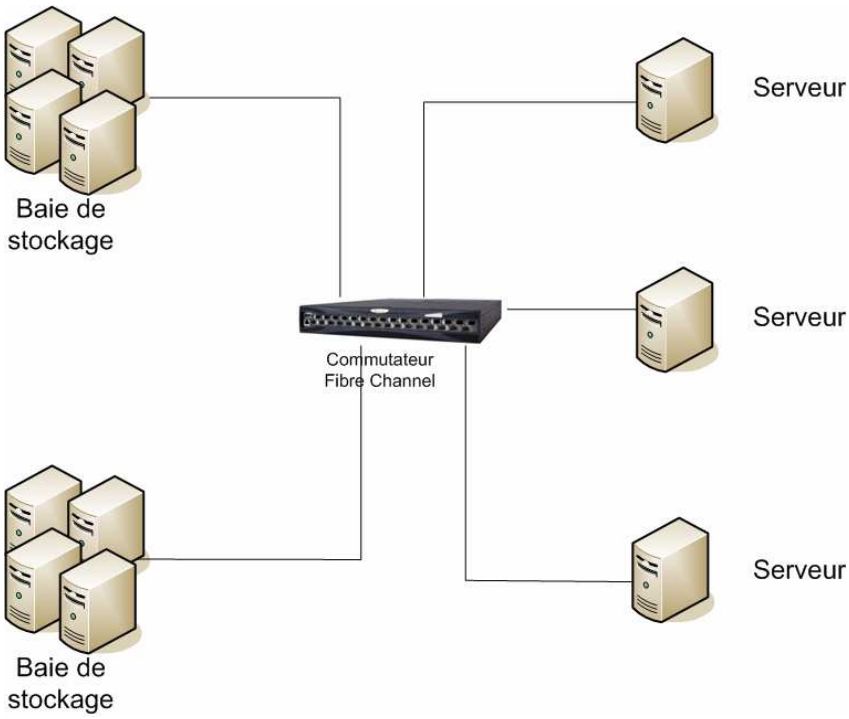

Figure 11: Topologie en étoile

- La topologie « en maille », appelée aussi Fabric apporte le meilleur taux de disponibilité car il est possible de doubler chaque lien. De plus chaque communication possède sa propre bande passante contrairement à la topologie en boucle. C'est aussi la topologie qui demande le plus grand investissement à cause de l'achat d'éléments réseaux spécifiques.

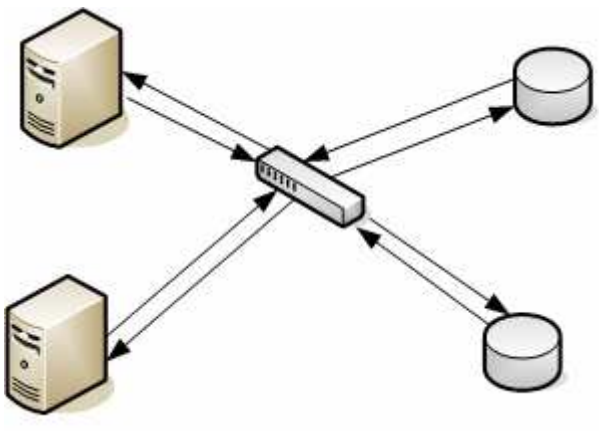

Figure 12: Topologie en maille

5.3.5 Avantages et inconvénients

Avantages

Performance

La performance du SAN est directement liée à celle du type de réseau utilisé. Dans le cas d'un réseau Fibre Channel, la bande passante est d'environ 100 Mo/s (800 Mbit/s) et peut être étendue en multipliant les liens d'accès.

Grâce au SAN, il est possible de partager des données entre plusieurs ordinateurs du réseau sans sacrifier les performances, dans la mesure où le trafic SAN est complètement séparé du trafic utilisateurs.

Ce sont les serveurs applicatifs qui jouent le rôle d'interface entre le réseau de données (généralement Fibre Channel) et le réseau des utilisateurs (généralement Ethernet).

Souplesse d'évolution

Le SAN permet une réelle souplesse d'évolution. Son architecture permet l'ajout d'espace de stockage, en ajoutant des disques Fibre Channel, sans avoir à rendre indisponible le reste des données. De ce fait, la capacité d'un SAN peut être étendue de manière quasi-illimitée et atteindre des centaines, voire des milliers de téraoctets.

Indépendance du stockage

Le SAN permet ainsi de dissocier le stockage des serveurs d'applications. Mais aussi, l'administration des données est ainsi facilitée, car centralisée.

La virtualisation du stockage

Une allocation d'un espace de stockage est destinée à chaque serveur. Mais le serveur ne sait pas exactement comment se présente physiquement le stockage. Son espace de stockage correspond à une partie de l'espace virtuel disponible.

Cette virtualisation permet une réorganisation dynamique en fonction du besoin de l'entreprise. Tout ce qui concerne le stockage n'est plus du tout géré par leur serveur d'application, mais par le SAN (*cf.* **Figure 13: Abstraction des éléments physiques pour les serveurs d'applications**, *page 41*).

Inconvénients

En contrepartie, le coût d'acquisition d'un SAN est beaucoup plus onéreux qu'un dispositif NAS dans la mesure où il s'agit d'une architecture complète, utilisant des technologies encore chères.

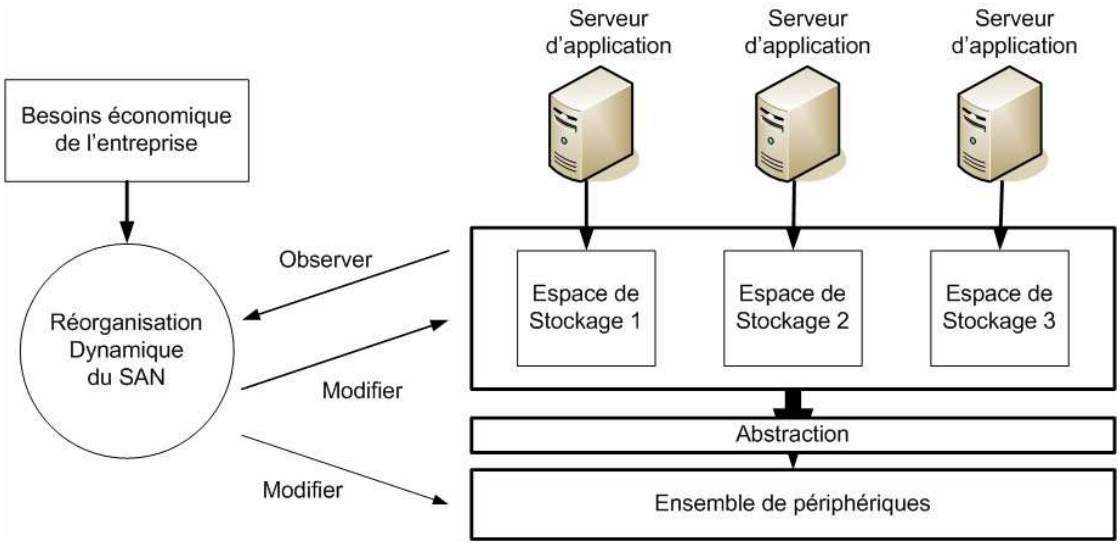

Figure 13: Abstraction des éléments physiques pour les serveurs d'applications

5.4 SAN sur IP: iSCSI

Les besoins de stockage de données ne cessent de croître. Le mariage du protocole Internet (IP) et du SCSI (Small Computer System Interface), le protocole le plus répandu pour le stockage pour créer une nouvelle solution de stockage réseau est donc une bonne nouvelle. L'émergence de iSCSI devrait ainsi fournir aux entreprises une solution de stockage réseau supérieure par rapport aux solutions de stockage suscités.

5.4.1 Principe de fonctionnement

Le iSCSI (Internet Small Computer System Interface) est un hybride qui rend compatibles des protocoles qui ne le sont pas naturellement : SCSI (le protocole le plus utilisé pour le stockage) et TCP/IP, en encapsulant les données SCSI dans des paquets IP.

iSCSI est donc un protocole qui autorise le transfert de blocs de données sur des réseaux Ethernet via des adresses IP.

À la différence de la solution NAS (Network Attached Storage), qui propose un accès aux données sous la forme de fichiers, iSCSI propose un accès au niveau des blocs de données. L'avantage serait donc des performances très supérieures.

Avant l'introduction de iSCSI, le seul moyen de transférer des données réseau consistait à utiliser les formats E/S des fichiers. L'introduction de données de stockage de bloc E/S améliore les performances, car elle supprime le besoin de traduire vers des formats de fichiers E/S. iSCSI est donc une solution de SAN sur réseau IP.

5.4.2 Modèle du protocole iSCSI

Lorsqu'un utilisateur final ou une application envoie une requête, le système d'exploitation génère la commande SCSI appropriée et la requête de données, qui est ensuite encapsulée et, si nécessaire cryptée. Un entête de paquet est ajouté avant que le paquet IP résultant soit transmis sur la connexion ethernet.

Lorsqu'un paquet est reçu, il est décrypté (s'il a été crypté auparavant) et est décomposé, en séparant la commande SCSI du paquet IP.

Le paquet SCSI composé de l'entête iSCSI et les données SCSI sont envoyées au contrôleur iSCSI. L'entête iSCSI est utilisé par ce contrôleur pour extraire et stocker les blocks de données d'E/S. Le contrôleur retransmet ensuite au matériel de stockage SCSI les données et les commandes SCSI.

Comme iSCSI est bidirectionnel, le protocole peut aussi être utilisé pour retourner des données dans la réponse à la requête originale.

iSCSI surveille les transferts en mode bloc et valide les opérations de lecture-écriture, au travers d'une ou plusieurs connexions TCP entre cible et émetteur. Ainsi, ces informations traversent toutes la couche comme indiquée sur la figure ci-après :

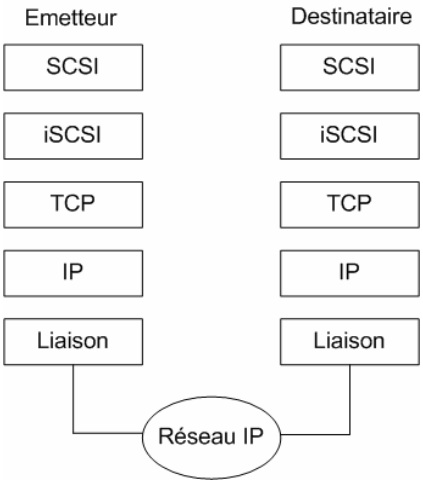

Figure 14: Modèle du protocole iSCSI

5.4.3 Topologie

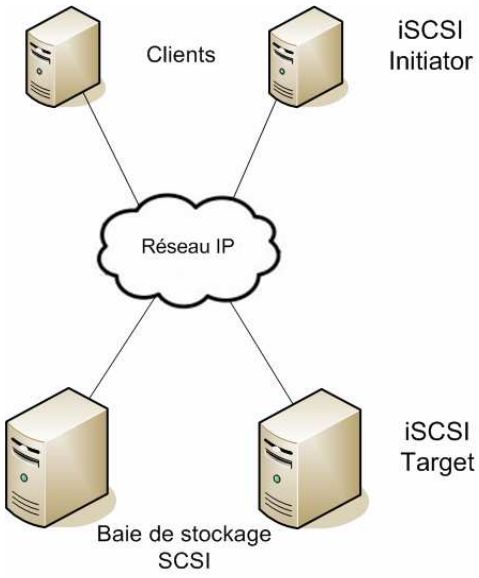

Figure 15: Exemple de topologie iSCSI

5.4.4 Avantages et inconvénients

Avantages

L'un des principaux atouts de iSCSI sur les technologies concurrentes est de tirer avantage de l'infrastructure IP existante. Cela se traduit par une réduction des coûts sur plusieurs points :

- La mise à jour du matériel,
- La formation des utilisateurs,
- La rapidité du déploiement.

On peut également ajouter d'autres points forts :

- La fiabilité : celle de l'infrastructure IP sur laquelle repose iSCSI.
- Un stockage agrégé : à la différence de certaines solutions qui isolent des zones de stockage, iSCSI crée un stockage partagé. Les utilisateurs peuvent ainsi se connecter à plusieurs périphériques sur plusieurs serveurs pour disposer de plus d'espaces de stockage et gérer plus facilement les données.
- La performance : iSCSI améliore la performance en offrant aux données un accès au niveau des blocs, plutôt qu'au niveau des fichiers. La performance est également améliorée du fait qu'une partie du traitement SCSI est gérée par la carte réseau.

Inconvénients

Les débits qu'offre cette solution sont inférieurs à ceux de la fibre optique, puisqu'ils sont indexés au débit des réseaux Ethernet, soit 10/100Mbps (contre 1 voire 2 Gbps pour certains équipements Fibre Channel).

5.5 Comparatif NAS/SAN

5.5.1 Différences entre NAS et SAN

SAN	NAS
Orienté paquets SCSI	Orienté fichier
Basé sur le protocole Fibre Channel	Basé sur le protocole Ethernet
Le stockage est isolé et protégé de l'accès client général	Conçu spécifiquement pour un accès client général
Support des applications serveur avec haut niveau de performances SCSI	Support des applications client dans un environnement NFS/CIFS hétérogène
Le déploiement est souvent complexe	Peut être installé rapidement et facilement

Tableau 1: Différences entre NAS et SAN

5.5.2 Vue générale des technologies NAS et SAN

	SAN	NAS
Fonction principale	Le stockage est accessible à travers un réseau qui lui est spécialement dédié. Sa principale fonction est de fournir aux serveurs un stockage consolidé basé sur le Fibre Channel	Serveur spécialisé, qui sert les fichiers et les données stockées aux postes clients et aux autres serveurs à travers le réseau

Applications bien adaptées	Idéal pour les bases de données et le traitement des transactions en ligne	Idéal pour serveur de fichiers
Transfert des données	A travers le SAN vers un serveur, vers un LAN ou un WAN	A travers un LAN ou un WAN
Ressources de stockage et de sauvegarde	Les ressources de stockage et de sauvegarde peuvent être attachées directement au serveur ou à travers une structure Fibre Channel	Les sauvegardes peuvent être attachées directement à des appliances NAS intermédiaires ou être distribuées et attachées à un LAN ou un WAN
Disponibilité	Des composants matériels et logiciels redondants donnent au système une haute disponibilité. Le système peut être configuré sans le moindre point de panne	Des alimentations et des ventilateurs redondants sont couramment utilisés
Scalabilité	Le stockage peut être étendu par l'ajout de switch Fibre Channel et de dispositifs de stockage	Plusieurs serveurs NAS peuvent être ajoutés au réseau, et du stockage peut être ajouté aux serveurs NAS intermédiaires

Tableau 2: Vue générale des technologies NAS et SAN

5.6 Cohabitation NAS et SAN

Le NAS et le SAN sont deux solutions de stockage différentes, et de par leurs caractéristiques différentes, les combiner peut permettre aux entreprises de mettre à profit ces différences.

5.6.1 Pourquoi?

On peut supposer que la convergence NAS/SAN est une réalité qui va s'affirmer dans l'avenir comme une composante importante du traitement et de la circulation des données au sein des grandes et moyennes entreprises.

Les utilisateurs accèderont aux données via des disques NAS sécurisés et reliés au SAN, réservé aux applications stratégiques, à la consolidation et à la sauvegarde de toutes les données d'entreprise sur place et sur des sites distants.

Grâce à cette combinaison, les réseaux d'entreprise pourront par exemple:

- Appliquer un degré de sécurité sur les données,
- Assurer une continuité de service auprès de leurs collaborateurs, clients et fournisseurs,
- Augmenter les performances de certaines applications,
- Améliorer la réactivité,
- Etre plus ouverts à l'évolution,
- Optimiser le trafic,
- Simplifier l'administration et la maintenance,
- Centraliser les alarmes.

5.6.2 Architecture

Le schéma suivant montre le principe de cohabitation de NAS et SAN. On peut remarquer que le NAS est toujours directement relié au réseau de l'entreprise, et le SAN est toujours un réseau à part connecté au réseau de l'entreprise.

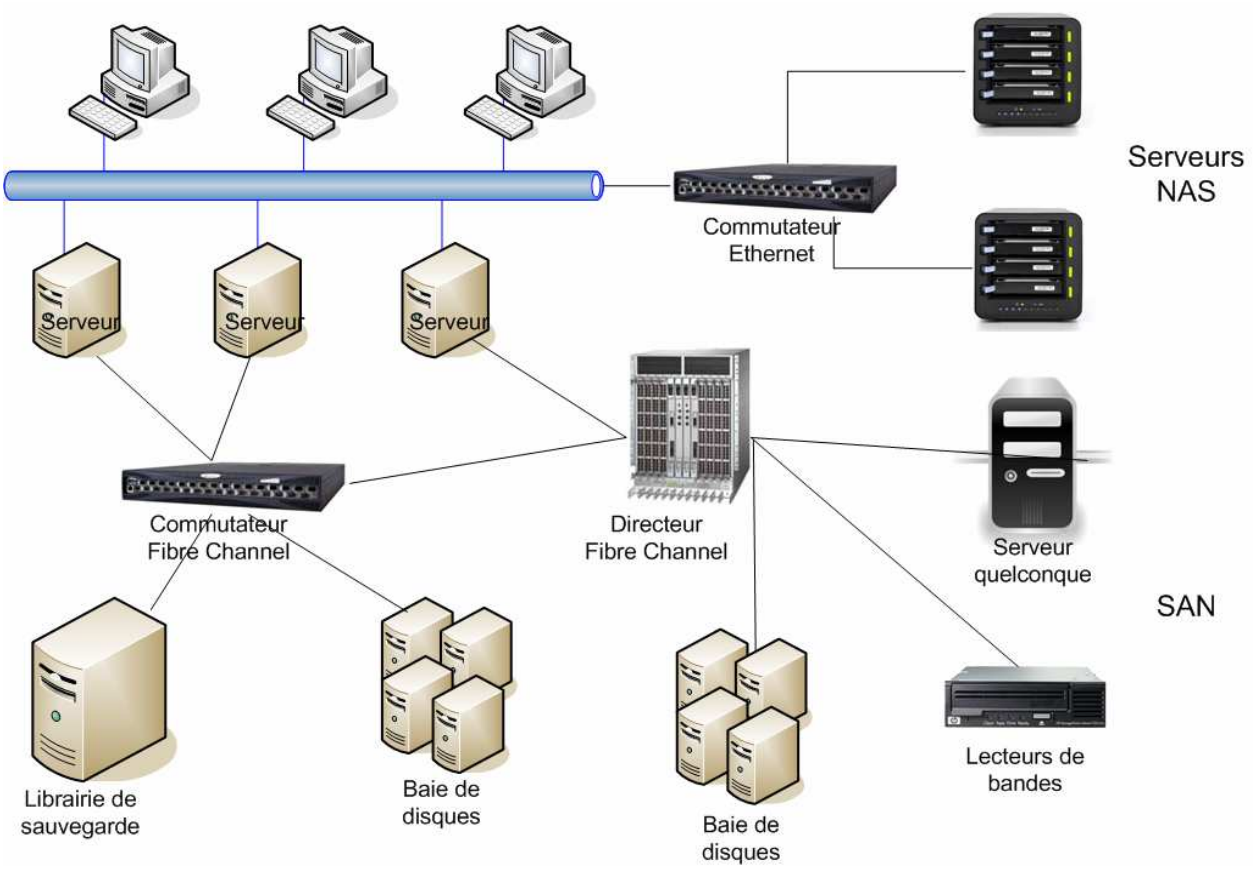

Figure 16: Cohabitation NAS et SAN

5.7 Que choisir ?

Le SAN et le NAS permettent l'obtention d'un stockage de données sur le réseau, mais leur philosophie différente impose une étude des besoins avant de choisir l'une des deux solutions.

5.7.1 Critères de choix

Avant de faire un choix, il faut connaître les différents critères intervenant dans le choix d'un NAS ou d'un SAN:

- Compatibilité OS,
- Volume de stockage,
- Administration,
- Installation,
- QoS,
- Coût,
- Disponibilité.

La question qui se pose est : quelle technologie choisir en fonction de ces critères ?

Compatibilité OS

Le NAS convient bien aux environnements hétérogènes. En effet, l'interopérabilité des éléments de stockage ne dépend que de l'OS du serveur de fichiers. En effet, un OS Microsoft et un Unix n'utilisent pas le même protocole de partage de fichiers :

- NFS (Network File System) pour Unix,
- CIFS (Common Internet File System) pour Windows.

Le SAN multipliant le matériel de stockage nécessaire (serveurs, mais aussi commutateurs ou routeurs et baies de disques issues de constructeurs différents), l'interopérabilité devient compliquée.

Volume de stockage

Que ce soit SAN ou NAS, les deux solutions de stockage permettent de supporter une grande quantité de stockage et de la faire évoluer.

Toutefois, il faut préférer un réseau SAN pour les réseaux connaissant une croissance rapide, ou qui ont besoin d'augmenter leurs capacités de stockage de façon sporadique.

Administration

Les serveurs NAS sont les plus simples à gérer dans la mesure où ils s'intègrent directement au réseau de l'entreprise. Et les réseaux SAN, grâce à des logiciels de configuration, simplifient grandement l'administration du stockage.

Installation

On peut ajouter des serveurs NAS au réseau local en quelques minutes. Ces serveurs sont particulièrement adaptés aux applications qui impliquent de nombreux accès en lecture/écriture.

SAN par contre est un réseau à part, il faut donc étudier tout ce réseau avant de pouvoir le mettre à disposition des utilisateurs.

QoS

Ici, le NAS montre ses faiblesses. En effet, le réseau Ethernet sur lequel repose le NAS n'offre en aucune garantie quant au fait que la requête envoyée par un serveur a été bien reçue et prise en compte par le système de stockage.

Alors que pour SAN, le commutateur prend en charge cette fonction et garantit en outre un débit fixe (100Mo/s par lien en fibre optique).

Ainsi, les entreprises ayant des applications critiques nécessitant une haute qualité de service devront opter pour le SAN.

Coût

Le NAS ne requiert pas de l'entreprise qu'elle mette en place une infrastructure de câbles en fibre optique (solution majoritairement adoptée pour le SAN). Son prix est donc abordable pour des petites entreprises ou des services départementaux de grands groupes dont les volumes de données ne sont pas trop importants.

C'est en dernière instance un arbitrage entre coût et besoins qui doit décider de l'intérêt d'une solution ou d'une autre.

Disponibilité

Le SAN assure la redondance du stockage en doublant au minimum chacun des éléments du système: les cartes HBA (Host Bus Adapter) des serveurs, les commutateurs, et l'écriture des données sur les disques.

Le NAS lui ne permet pas cette fonction vitale pour certaines applications (type bancaires, assurances, sites de commerce électronique…).

5.8 Quel stockage pour quelle application ?

Il est nécessaire de préciser quel stockage il faut utiliser pour une application spécifique.

Le NAS est spécifique pour les serveurs de fichiers, c'est-à-dire le partage de fichiers. Il est utilisé pour stocker les données non critiques mais sauvegardées régulièrement et pour le travail collaboratif.

Le SAN, de son côté, est destiné aux applications ne supportant pas le mode fichiers et écrites pour une utilisation sur un serveur local comme le SQL Server, Exchange,…. Il est utilisé pour les applications hautement transactionnelles et pour le stockage et la sauvegarde des données critiques.

5.9 Résumé

En tenant compte des faits que :

- la société Blueline veut mettre en place un système de stockage distribué et si possible pouvant fonctionner avec le système openvz,
- le système de stockage par NFS ne fonctionne pas avec openvz et
- les avantages du SAN (l'ajout d'espace de stockage sans aucune interruption de service, une grande performance, une facilité d'administration et surtout la virtualisation du stockage).

Il est clair que c'est la solution SAN qui convient le plus.

Or, un des inconvénients du SAN est son coût. Le iSCSI est une alternative puisqu'il permet de réaliser un réseau de stockage à moindre coût, tout en profitant de l'infrastructure et de l'expertise IP déjà existante.

Chapitre 6

Architecture réseau de la solution retenue

A part l'iSCSI, il existe d'autres technologies utiles pour la mise en place d'un système de stockage distribué. Ce chapitre présente l'architecture réseau de la solution retenue et explique une à une les technologies utilisées.

6.1 Architecture réseau de la solution retenue

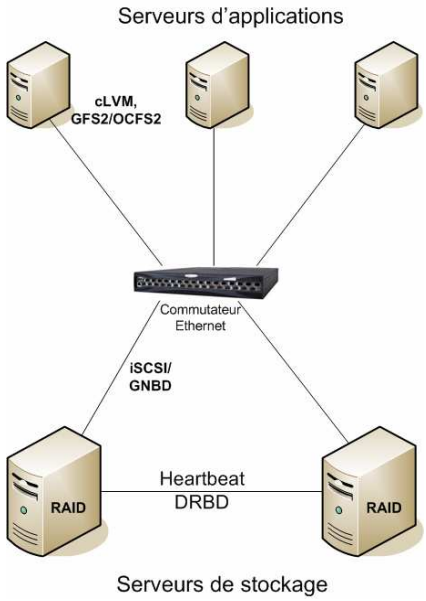

Figure 17: Architecture réseau de la solution retenue

Les serveurs d'applications peuvent être : serveur mail, serveur web, serveur ftp, etc. Le stockage des données est centralisé sur les serveurs de stockage.

L'utilisation de la technologie RAID permet d'avoir une redondance des données en locale.

Pour ce faire, on a besoin d':

- Exporter les disques depuis les serveurs de stockage vers les serveurs d'applications,
- Utiliser un gestionnaire de volumes et
- Utiliser un système de fichier en cluster.

Remarque :

Chaque serveur de stockage est équipé d'un système de réplication, au cas de panne de l'un, l'autre prend le relais. Le couple Heartbeat et DRBD effectuent cette tâche. L'étude et la mise en place de cette réplication a fait l'objet d'un autre projet et ne sera pas détaillé dans ce qui suit.

6.2 Exporter des disques en réseau

6.2.1 iSCSI

iSCSI est un protocole réseau qui encapsule le protocole SCSI dans des paquets TCP. TCP/IP est alors utilisé comme protocole de transport, ce qui permet à une machine d'accéder à des périphériques distants connectés à l'infrastructure réseau existante.

Le protocole iSCSI utilise l'architecture client/serveur, mais il possède sa propre terminologie:

- le client est appelé « initiator » et
- le serveur est appelé « target ».

(cf. **Figure 18: Architecture iSCSI versus FC/SCSI,** *page 52*).

L'initiator peut être une carte Ethernet spécialisée, munie d'un composant supplémentaire qui va gérer le protocole iSCSI. Cette carte sera appelée un " iSCSI-HBA ". Mais, on peut aussi déployer une solution logicielle pure qui nécessite seulement un driver spécifique dont le rôle sera de traduire les ordres SCSI en paquets réseau.

Le target (cible) est normalement un périphérique de stockage doté d'une interface iSCSI. Initiator et target sont considérés comme les nœuds d'un réseau iSCSI.

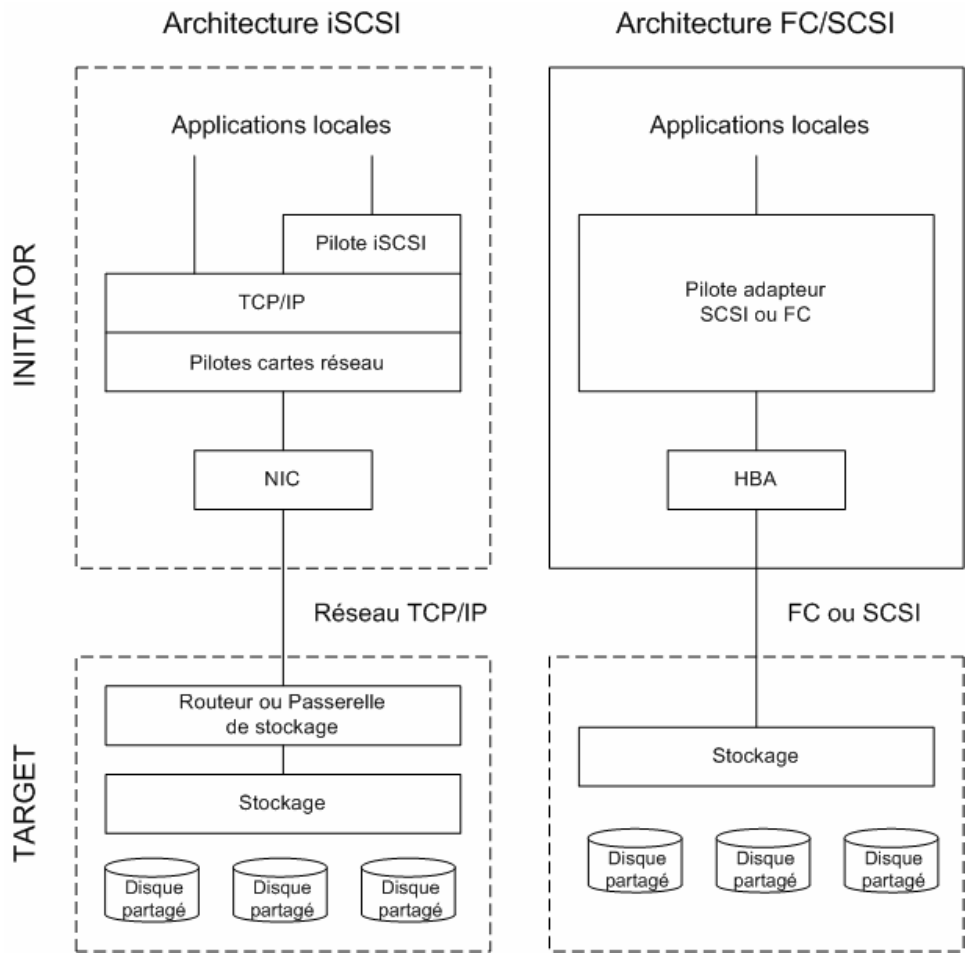

Figure 18: Architecture iSCSI versus FC/SCSI

Remarque :

L'espace de stockage étant accédé en mode bloc à travers le réseau. Le noyau Linux considère cet espace disque comme s'il était local. Il n'est pas question de partager le même espace physique entre plusieurs initiators (à moins d'utiliser un système de fichiers distribué comme GFS et OCFS).

On appliquera alors les mêmes règles qu'avec un SAN : si vous avez « n » initiators, alors il faut découper l'espace physique en « n » portions logiques pour faire en sorte que chaque initiator accède à son espace logique.

6.2.2 Périphérique bloc du réseau global (GNBD)

GNBD fournit un accès aux périphériques blocs à Red Hat GFS à travers TCP/IP. GNBD est similaire dans son concept à NBD ; cependant, GNBD est spécifique à GFS et personnalisé pour être utilisé exclusivement avec GFS.

GNBD est composé de deux principaux éléments: un client et un serveur.

Le GNBD serveur exporte un périphérique qui lui est local et le GNBD client l'importe et l'utilise comme s'il était un disque local.

Plusieurs GNBD clients peuvent accéder à un périphérique exporté par un GNBD serveur. GNBD peut donc être utilisé dans un environnement en cluster.

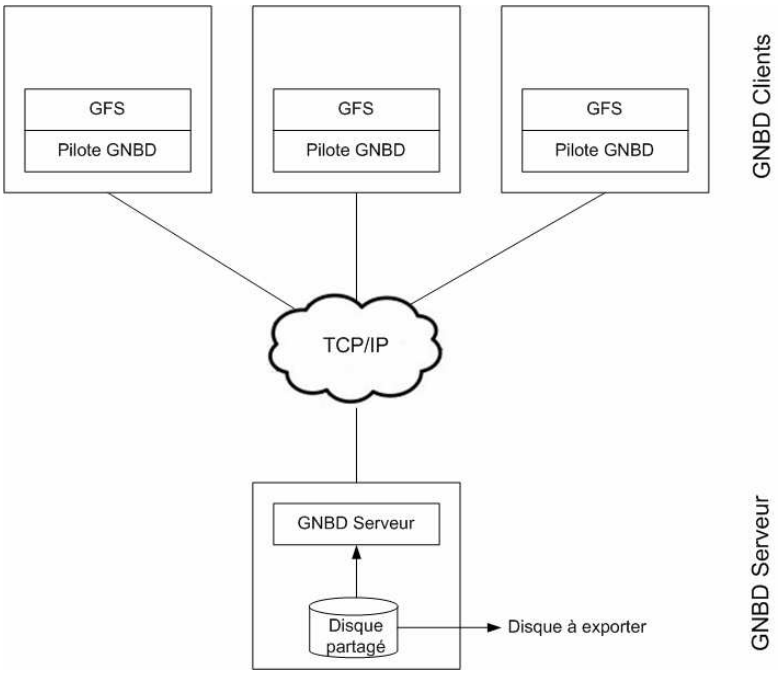

Figure 19: Aperçu de GNBD

6.3 Systèmes de fichiers en cluster

Les systèmes de fichiers partagés permettent le partage de données par le réseau. Le principe est simple, les données se trouvent sur un serveur ne faisant pas partie du cluster et les nœuds du cluster peuvent consulter ces mêmes données :

6.3.1 GFS

Avantages

GFS est un système de fichiers prévu pour être utilisé sur des disques partagés (c'est-à-dire reliés physiquement à plusieurs serveurs) sous Linux. Les disques peuvent être accédés simultanément en lecture et en écriture depuis tous les serveurs qui leur sont connectés. GFS fournit en plus une fonction de journalisation.

Architecture

Le schéma ci-dessous présente l'architecture requise pour utiliser GFS :

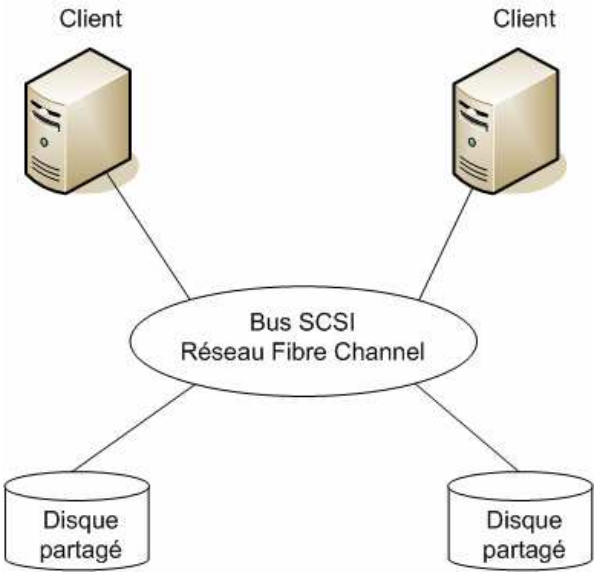

Figure 20: Architecture GFS

GFS supporte un nombre quelconque de clients reliés à un ou plusieurs disques par un moyen de communication pouvant être l'un des suivants :

- Un bus SCSI ;
- Un réseau Fibre Channel ;
- Un réseau IP.

Dans ce dernier cas, il faut adopter une solution permettant d'utiliser un disque distant par l'intermédiaire d'un réseau IP. Les solutions possibles comprennent NBD (Network Block Device) livré en standard avec Linux, ou GNBD.

GFS permet d'accéder aux disques simples ou aux baies RAID.

GFS se présente sous la forme d'un module ajoutant le support du système de fichiers GFS et d'utilitaires permettant la création et la gestion de disques ou de partitions partagés.

Caractéristiques

Le Global File System est un système de fichiers journalisé qui permet de manipuler sans risque des données présentes sur un support de stockage partagé par plusieurs machines.

Sa caractéristique principale est de laisser les systèmes lire et écrire directement les données. Celles-ci ne passent pas par le réseau (ce qui permet un gain important au niveau des performances par rapport à un protocole comme NFS), seul transite un flux protocolaire permettant de verrouiller les accès et d'éviter les conflits. GFS coordonne en effet les accès en écriture à travers un gestionnaire de verrous distribués (DLM pour Distributed Lock Manager) qui permet de garantir l'intégrité des données.

GFS supporte les baies de disques SCSI ou Fibre Channel ainsi que les protocoles iSCSI.

La plus grosse configuration basée sur GFS a été présentée à la Linux World Expo 2000, et comprenait 16 clients et 64 disques répartis en 8 baies, le tout relié par Fibre Channel.

6.3.2 OCFS

OCFS ou Oracle Cluster File System, comme GFS, est un système de fichier en cluster. Il a été développé par Oracle et a été intégré au noyau Linux depuis la version 2.6.16.

OCFS utilise un DL M (Distributed Lock Manager) pour verrouiller les fichiers en fonction des nœuds qui les demandent.

6.4 Gestionnaire de volumes logiques

6.4.1 LVM

LVM ou Logical Volume Manager, est un gestionnaire de volumes logiques qui apporte une vision abstraite du stockage. La partition physique devient un composant élémentaire et ne va plus directement supporter le système de fichiers.

Entre ces deux couches, successivement la partition et le système de fichiers, il existe :

- Un volume groupe : une première couche regroupant les partitions physiques, et
- Des volumes logiques : une deuxième couche dans laquelle s'effectue la création d'un système de fichiers. Les volumes logiques peuvent être redimensionner à chaud, indifféremment en réduction ou en extension.

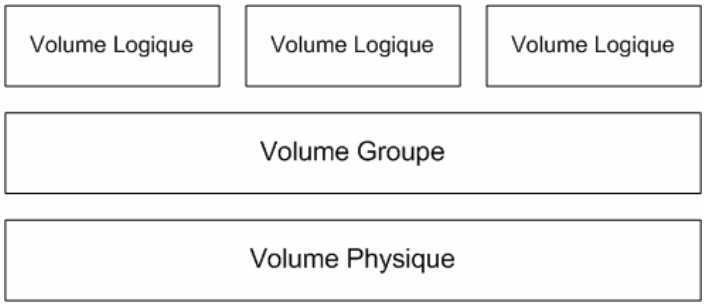

Figure 21: Découpage des données sous LVM

L'intérêt principal de cette approche est à la fois de pouvoir étendre facilement un volume groupe en ajoutant des partitions physiques (par exemple, un nouveau disque dur) et de changer la taille à chaud des volumes logiques. Voici un exemple de ce que pourrait donner un système tournant sur LVM (*cf.* **Figure 22: Couches d'abstractions du modèle LVM**, *page 56*).

Le nommage des volumes groupes et logiques est libre même si la convention habituelle est de préfixer un volume groupe par « vg_ » et un volume logique par « lv_ ». Tout l'espace disponible dans un volume groupe peut ne pas être utilisé dans les volumes logiques. Cela permet de garder en réserve de l'espace disque pour l'ajouter aux volumes logiques en cas de besoin.

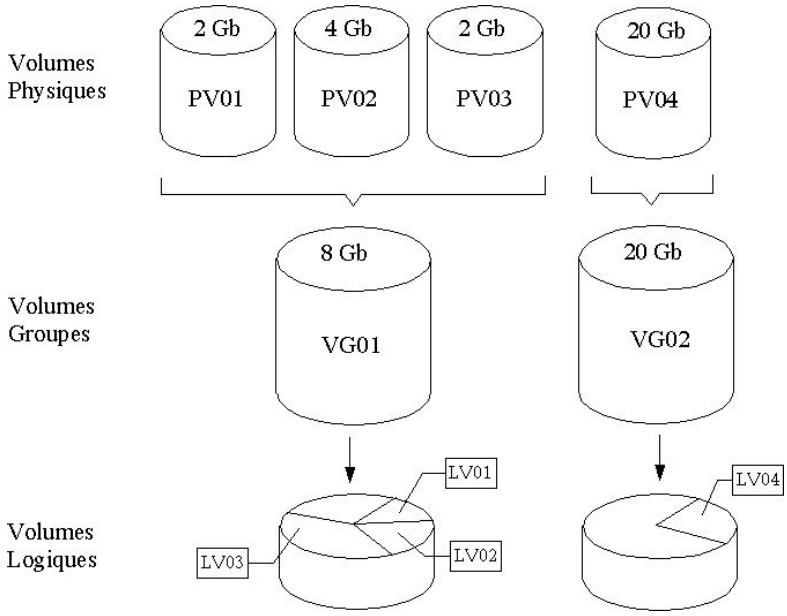

Figure 22: Couches d'abstractions du modèle LVM

Dans LVM, un volume groupe est divisé en volumes logiques. Il a trois types de volumes logiques : un volume linéaire, un volume strippé, et un volume mirroré.

Un volume linéaire

Un volume linéaire agrège plusieurs volumes physiques en un seul volume logique. Par exemple, si on a deux disques de 60 GB, on peut créer un disque logique de 120 GB. Les volumes physiques sont donc concaténés.

Remarque :

Chaque volume physique est divisé en morceaux de données, appelés extents physiques. Ces extents ont une taille identique à celle des extents logiques du groupe de volumes.

Chaque volume logique est divisé en morceaux de données, appelés extents logiques. La taille d'extents est la même pour tous les volumes logiques du groupe de volumes.

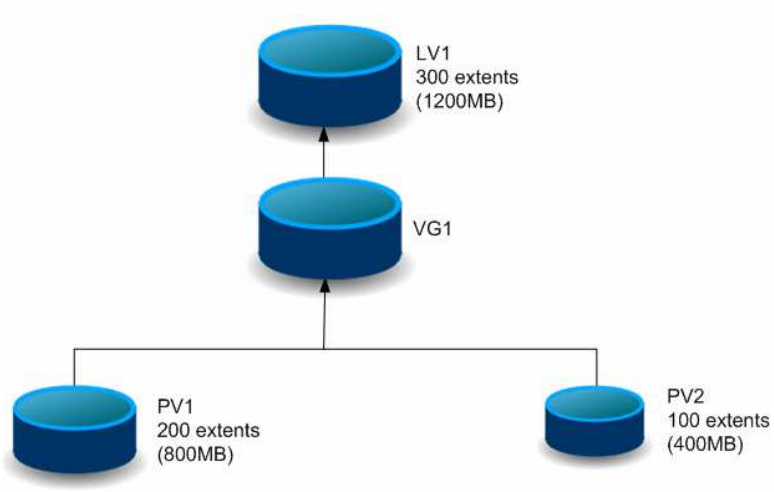

Figure 23: Exemple d'un volume linéaire

Un volume strippé

Lors de l'écriture des données sur un volume logique, le système répartit les donnés dans l'ensemble des disques physiques. Dans ce cas, l'écriture sur les volumes physiques peut être effectuée en créant un volume logique strippé.

Pour les E/S importantes, cela peut améliorer les performances. Le stripping augmente les performances en écrivant les données dans un nombre prédéterminé de disques physiques en mode série. Avec le stripping, les E/S peuvent être fait en parallèle. Dans certaines situations, cela peut entraîner une performance linéaire pour chaque disque physique installé.

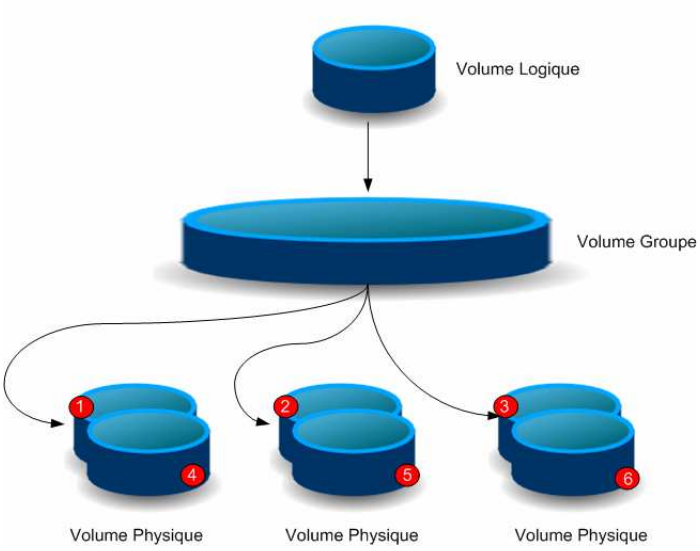

Figure 24: Cas d'un volume strippé

Un volume mirroré

Un miroir maintient une copie identique des données contenues sur différents périphériques. Quand une donnée est écrite sur un des périphériques, elle est automatiquement écrite sur le deuxième. Cela protège des pannes matérielles. Quand une des deux parties du mirroir cède, le volume devient alors linéaire, et reste accessible.

LVM supporte les volumes mirrorés. Lors de la création d'un volume mirroré, LVM s'assure que les données écrites sur un des volumes physiques soient bien répliquées sur un volume physique distinct. Avec LVM, on peut créer un volume logique avec plusieurs miroirs.

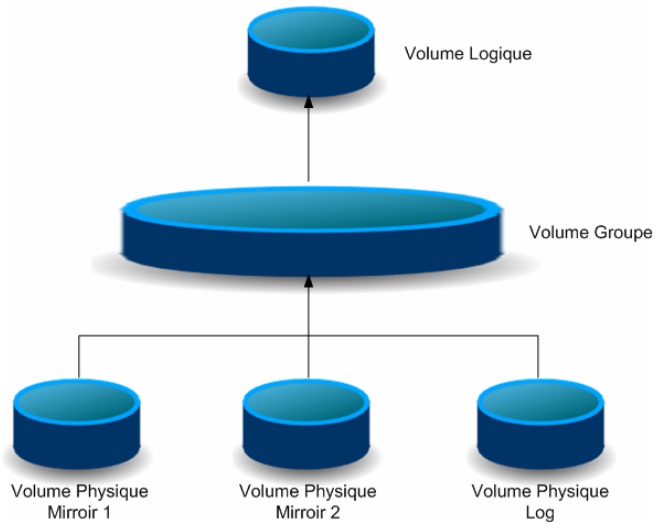

Figure 25: Cas d'un volume mirroré avec log sur disque

6.4.2 CLVM

Le Clustered Logical Volume Manager (CLVM) est un ensemble d'extensions de cluster de LVM. Ces extensions permettent à un cluster de machine de gérer un stockage partagé (par exemple un SAN) tout en utilisant LVM.

Le daemon clmvd est l'outil qui permet le fonctionnement de LVM en cluster. Le daemon clvmd tourne sur chaque machine constituant le cluster. Il distribue les mises à jours des « metadatas » LVM au sein du cluster et présente chaque machine faisant parti du cluster avec la même vue du volume logique.

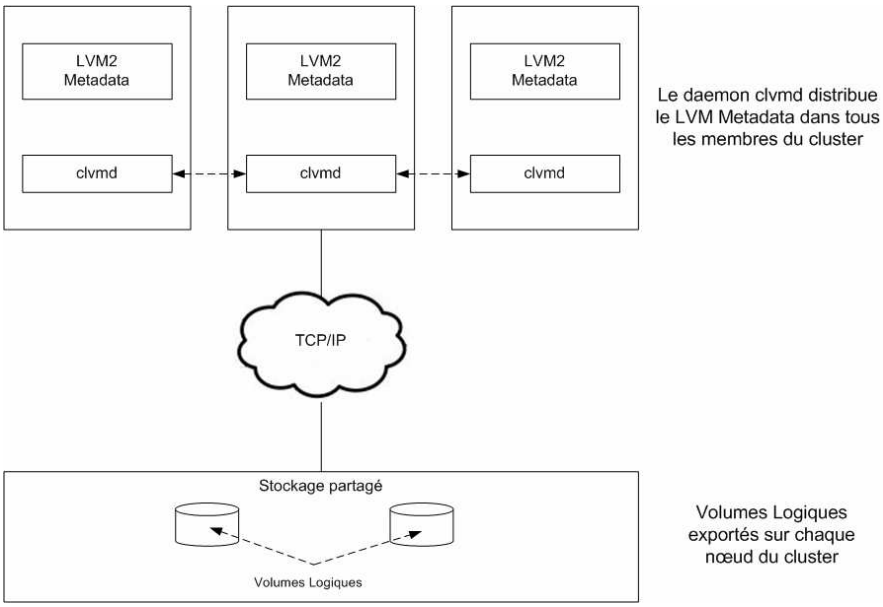

Figure 26: Vue d'une architecture CLVM

Les volumes logiques créés avec CLVM sur un stockage partagé sont visibles sur toutes les machines qui ont accès à ce stockage partagé.

CLVM permet à un utilisateur de configurer des volumes logiques sur un stockage partagé en verrouillant l'accès au stockage physique pendant que le volume logique est en train d'être configuré.

Partie IV

Réalisation

CHAPITRE 7 :

Installation et utilisation de OCFS2 et GFS2

CHAPITRE 8 :

Mise en œuvre de NFS, iSCSI et GNBD et test de performance

CHAPITRE 9 :

Gestion du système de fichiers avec LVM

CHAPITRE 10 :

Trio iSCSI, CLVM et OCFS2

Chapitre 7

Installation et utilisation de OCFS2 et GFS2

7.1 OCFS2

7.1.1 Installation

```
# apt-get install ocfs2-tools
```

7.1.2 Configuration

Le fichier /etc/cluster/cluster.conf

Le système de fichier OCFS2 est déstiné â être utiliser dans un environnement en cluster. Son fichier de configuration est **/etc/ocfs2/cluster.conf**. Dans ce fichier, il faut spécifier tous les noeuds du cluster. Il doit être absolument identique sur tous les noeuds du cluster.

Un exemple du fichier **/etc/ocfs2/cluster.conf** :

```
cluster:
      node_count = 2
      name = racdb
node:
      ip_port = 7777
      ip_address = 192.168.1.110
      number = 0
      name = node1
      cluster = racdb
node:
      ip_port = 7777
      ip_address = 192.168.1.111
      number = 1
      name = node2
      cluster = racdb
```

Le mot clé **cluster** a deux paramètres:

- *node_count:* le nombre total des noeuds du cluster,
- *name:* le nom du cluster.

Le mot clé **node** a cinq paramètres:

- *ip_port:* le numéro du port utilisé,
- *ip_address:* l'adresse IP de la machine,
- *number: le* numéro du noeud compris entre 0 et 254,
- *name: le* nom de la machine (hostname),
- *cluster: le* nom du cluster qui est le même que dans le mot clé cluster.

Démarrer OCFS2 au démarrage

Pour activer **o2cb** lors du démarrage du système, éditer le fichier **/etc/default/o2cb** et changer la valeur de **O2CB_ENABLED** à **true** :

```
# nano /etc/default/o2cb
...
# O2CB_ENABLED: 'true' means to load the driver on boot.
O2CB_ENABLED=true
....
```

Pour démarrer et stopper le cluster **racdb** :

```
# /etc/init.d/o2cb start|stop
```

7.1.3 Tâches courantes

Formater une partition avec OCFS2

```
# mkfs.ocfs2 –L racdb /dev/sda1
```

Monter une partition OCFS2

```
# mkdir /mnt/racdb
# mount.ocfs2 /dev/sda1 /mnt/racdb
```

Agrandir une partition OCFS2

```
# tunefs.ocfs2 –S /dev/sda1
```

La commande **tunefs.ocfs2** permet de changer les paramètres d'un système de fichiers OCFS2.

tunefs.ocfs2 -S *BlockDevice*

-S

Agrandir la taille du système de fichiers

BlockDevice

Précise un volume physique ou logique.

Remarque :

Pour plus d'informations sur l'utilisation de **mkfs.ocfs2 et tunefs.ocfs2**, il faut regarder la page du manuel.

```
# man mkfs.ocfs2
# man tunefs.ocfs2
```

7.2 GFS2

7.2.1 Installation

```
# apt-get install gfs2-tools
```

7.2.2 Configuration

Le fichier /etc/cluster/cluster.conf

Le système de fichiers GFS2 est un système de fichiers en cluster mais il peut être utilisé comme un système de fichiers local.

Lors d'une utilisation en cluster, il faut configurer le fichier **/etc/cluster/cluster.conf**.

Un exemple du fichier **/etc/cluster/cluster.conf** :

```
<?xml version="1.0"?>
<cluster name="racdb" config_version="1">

<cman two_node="1" expected_votes="1"/>

<clusternodes>
<clusternode name="node1" votes="1" nodeid="1">
        <fence>
                <method name="single">
                        <device name="manual" ipaddr="192.168.1.110"/>
                </method>
        </fence>
</clusternode>

<clusternode name="node2" votes="1" nodeid="2">
        <fence>
                <method name="single">
                        <device name="manual" ipaddr="192.168.1.111"/>
                </method>
        </fence>
</clusternode>
</clusternodes>

<fencedevices>
</fencedevices>

</cluster>
```

Le cluster **racdb** comporte deux nœuds respectivement **node1** et **node2**. Ce fichier de configuration est identique à celui utilisé par OCFS2, mais GFS2 utilise le format XML.

Le fichier /etc/lvm/lvm.conf

Editer le fichier /etc/lvm/lvm.conf et metter la valeur de **locking_type** à **2**.

```
# nano /etc/lvm/lvm.conf
locking_type = 2
```

7.2.3 Tâches courantes

Formater une partition avec GFS2

Utilisation

La création d'un système de fichiers GFS2 se fait à l'aide de la commande **mkfs.gfs2**. Le format de la commande est comme suit, respectivement pour l'utilisation en cluster et local :

```
mkfs.gfs2 -p LockProtoName -t LockTableName -j NumberJournals BlockDevice
mkfs.gfs2 -p LockProtoName -j NumberJournals BlockDevice
```

LockProtoName

Préciser le nom du protocole de verrouillage à utiliser. Le protocole de verrouillage à utiliser pour un cluster est **lock_dlm**, et en local **lock_nolock**.

LockTableName

Ce paramètre est précisé dans le système de fichiers GFS2 dans une configuration groupée. Ce paramètre est composé de deux parties séparées par deux points (sans espace) comme suit: ***ClusterName:FSName.***

ClusterName, le nom d'un cluster pour lequel le système de fichiers GFS2 a été créé.

FSName, le nom du système de fichiers, peut comporter de 1 à 16 caractères de long. Son nom doit être unique parmi les autres systèmes de fichiers lock_dlm présents dans le groupement, et pour tous les systèmes de fichiers (lock_dlm et lock_nolock) présents sur chaque noeud local.

NumberJournals

Spécifie le nombre de journaux que la commande mkfs.gfs2 doit créer. Un journal est utilisé par un nœud du cluster.

BlockDevice
Précise un volume physique ou logique.

Exemple

```
# mkfs.gfs2 -p lock_dlm -t alpha:mydata1 -j 8 /dev/vg01/lvol0
```

Monter une partition GFS2

Utilisation

mount *BlockDevice MountPoint*

BlockDevice
Précise le périphérique en mode bloc où le système de fichiers GFS2 se situe.

MountPoint
Précise le répertoire où le système de fichiers GFS2 devrait être monté.

Exemple

```
# mount /dev/vg01/lvol0 /mygfs2
```

Agrandir une partition GFS2

Utilisation

gfs2_grow *MountPoint*

MountPoint
Précise le système de fichiers GFS2 pour lequel les actions s'appliquent.

Exemple

Dans cet exemple, le système de fichiers de répertoire **/mnt/racdb** est agrandi.

```
# gfs2_grow /mnt/racdb
```

Chapitre 8

Mise en œuvre de NFS, iSCSI et GNBD et test de performance

8.1 Mise en œuvre de NFS

8.1.1 NFS Serveur

Installation des paquets

```
# apt-get install nfs-kernel-server nfs-common portmap
```

Configuration

Créer le répertoire à exporter :

```
# mkdir /data/export
```

Editer le fichier **/etc/exports** et ajouter la ligne suivante :

```
# nano /etc/exports
/data/export
192.168.1.0/24(rw,no_root_squash,no_all_squash,no_subtree_check,sync)
```

Redémarrer le service **nfs-kernel-service** :

```
# /etc/init.d/nfs-kernel-server restart
```

8.1.2 NFS Client

Installation des paquets

```
# apt-get install nfs-common portmap
```

Configuration

Créer le répertoire de montage :

```
# mkdir /data
```

Tester:

```
# mount 192.168.1.112:/data/export /data
# mount
...
192.168.1.112:/data/export on /data type nfs (rw,addr=192.168.1.112)
```

Editer le fichier **/etc/fstab** et ajouter la ligne suivante :

```
# nano /etc/fstab
192.168.1.112:/data/export /data nfs rw 0 0
```

8.2 Mise en œuvre de iSCSI

8.2.1 iSCSI Target

Installation des paquets

```
# apt-get install iscsitarget iscsitarget-modules-`uname -r`
```

Configuration

Pour lancer automatiquement le service lors du démarrage du système, ouvre le fichier **/etc/default/iscsitarget** et mettez la valeur de **ISCSITARGET_ENABLE** à **true**.

```
# nano /etc/default/iscsitarget
ISCSITARGET_ENABLE=true
```

Ensuite, éditer le fichier **/etc/ietd.conf** et commentez toutes les lignes. A la fin de ce fichier, ajouter les lignes suivantes :

```
#nano /etc/ietd.conf
Target iqn.2001-04.com.example:storage
      Lun 0 Path=/dev/sdXX,Type=blockio
      Alias STORAGE
```

Le nom suivi par le mot clé Target doit être unique, il est de la forme:

iqn.AAAA-MM.<nom du domaine a l'envers>[:identificateur]

Ce format est baptisé **iqn** (iSCSI Qualified Name).

AAAA-MM indique la date depuis laquelle le nom du domaine est valide.

identificateur est une chaîne unique qui identifie le noeud de stockage.

La ligne **Lun** spécifie le chemin complet du disque de stockage.

Remarque :

sdXX sera remplacé par le nom du disque à partager présent sur la machine Target.

Et pour terminer, il faut autoriser l'accès à la ressource de stockage pour tout le réseau local.

```
#nano /etc/initiators.allow
ALL 192.168.1.0
```

Maintenant, il faut tester le service pour s'assurer de son bon fonctionnement.

```
# /etc/init.d/iscsitarget restart
Removing iSCSI enterprise target devices: succeeded.
Stopping iSCSI enterprise target service: succeeded.
Removing iSCSI enterprise target modules: succeeded.
Starting iSCSI enterprise target service: succeeded.
#
```

Pour visualiser l'état de la session d'accès à la ressource de stockage iSCSI, il faut consulter le fichier **/proc/net/iet/session** qui donne l'état courant de la session avec l'adresse IP du système initiator.

```
# cat /proc/net/iet/session
tid:1 name:iqn.2001-04.com.example:storage
```

8.2.2 iSCSI Initiator

Installation des paquets

```
# apt-get install open-iscsi
```

Configuration

Pour rendre la connectivité à l'unité de stockage automatique lors du démarrage du système initiator, éditer le fichier **/etc/iscsi/iscsid.conf** et changer **manual** par **automatic**.

```
# nano /etc/iscsi/iscsid.conf
...
node.startup = automatic
...
```

Pour tester le bon fonctionnement du service, lancer la commande :

```
# /etc/init.d/open-iscsi restart
Disconnecting iSCSI targets:.
Stopping iSCSI initiator service:.
Starting iSCSI initiator service: iscsid
Setting up iSCSI targets:
iscsiadm: No records found!
.
Mounting network filesystems:.
#
```

Maintenant, on est en mesure de tester la connectivité iSCSI en utilisant la commande **iscsiadm**. Pour lister les disques de stockage présents sur la machine Target, lancez la commande :

```
# iscsiadm -m discovery -t st -p 192.168.1.112
192.168.0.100:3260,1 iqn.2001-04.com.example:storage
```

Et pour établir une connexion, utilisez la commande :

```
# iscsiadm -m node -T iqn.2001-04.com.example:storage -p 192.168.0.100 -l
Logging in to [iface: default, target: iqn.2001-04.com.example:storage,
portal:
192.168.0.112,3260]
Log in to [iface: default, target: iqn.2001-04.com.example:storage, portal:
192.168.0.112,3260]: successfull
```

Pour vérifier l'existence d'un nouveau disque, exécutez la commande suivante :

```
# fdisk -l
```

Partitionner et formater le nouveau disque :

```
# fdisk /dev/sdX
# mkfs.ocfs2 –L storage /dev/sdX1
# mkdir /mnt/storage
# mount /dev/sdX1 /mnt/storage
```

Editer le fichier **/etc/fstab** et ajouter la ligne suivante :

```
# nano /etc/fstab
/dev/sdX1 /mnt/storage ocfs2 rw 0 0
```

8.3 Mise en œuvre de GNBD

8.3.1 GNBD Serveur

Installation des paquets

```
# apt-get install gnbd-server
```

Configuration

Editer le fichier **/etc/cluster/gnbdexports.conf** et ajouter la ligne suivante :

```
# nano /etc/cluster/gnbdexports.conf
# <device> <exportname> <options>
/dev/sdX1    gnbddisk0  -c
```

La partition **/dev/sdX1** sera exportée sous le nom **gnbddisk0**.

Editer le fichier **/etc/default/gnbd-server** :

```
# nano /etc/default/gnbd-server
GNBD_OPTIONS="-n"
```

Remarque:

L'option "-n" indique qu'il ne faut pas rechercher un cluster. En cas d'utilisation de GNBD en cluster, il suffit de commenter cette ligne.

Démarrer le service **gnbd-server** :

```
# /etc/init.d/gnbd-server start
Starting global network block device server: gnbd_serv: startup succeeded
Exporting device(s):
gnbd_export: created GNBD gnbddisk0 serving file /dev/sdX1
done.
```

Lister les disques exportés:

```
# gnbd_export
Server[1] : gnbddisk0
--------------------------
      file : /dev/sdX1
   sectors : 204800
  readonly : no
    cached : yes
   timeout : no
```

8.3.2 GNBD Client

Installation des paquets

```
# apt-get install gnbd-client redhat-cluster-modules-`uname -r`
# modprobe gnbd
```

Configuration

Vérifier si le disque à importer existe :

```
# gnbd_import -e 192.168.1.112 -n
gnbddisk0
```

Importer le disque :

```
# gnbd_import -i 192.168.1.112 -n
gnbd_import: created directory /dev/gnbd
gnbd_import: created gnbd device gnbddisk0
gnbd_recvd: gnbd_recvd started
```

Lister le disque importé:

```
# gnbd_import -l -n
Device name : gnbddisk0
----------------------
    Minor # : 0
 sysfs name : /block/gnbd0
     Server : 192.168.1.112
       Port : 14567
      State : Close Connected Clear
   Readonly : No
    Sectors : 195318207
```

Formater et monter le disque importé :

Le disque importé porte le nom **/dev/gnbd/gnbddisk0** ou **/dev/gnbd0**.

```
# mkfs.gfs2 -p lock_nolock /dev/gnbd0
# mkdir /mnt/gnbd0
# mount /dev/gnbd0 /mnt/gnbd0
```

Editer le fichier **/etc/fstab** et ajouter la ligne suivante :

```
# nano /etc/fstab
/dev/gnbd0 /mnt/gnbd0 gfs2 rw 0 0
```

8.4 Test de performance du système de fichiers

Il faut effectuer un test de performance des systèmes de fichiers NFS, OCFS2 et GFS2 afin de connaître celui qui est le plus performant et le plus adapté aux besoins. OCFS2 sera utilisé avec iSCSI et GFS2 avec GNBD.

8.4.1 Mesurer la bande passante entre le client et le serveur

Pour effectuer le test, il faut deux machines : un client et un serveur.

Sur le serveur, lancer la commande :

```
# iperf -s -i 1
```

De même, sur le client, lancer la commande :

```
# iperf -s 1 –c 192.168.1.112
------------------------------------------------
Client connecting to stockagemaster, TCP port 5001
TCP windoz size: 16.0 KByte (default)
------------------------------------------------
[  3] local 192.168.1.110 port 48274 connected with 192.168.1.112 port 5001
[ ID] Interval       Transfer     Bandwidth
[  3]  0.0- 1.0 sec  12.7 MBytes    106 Mbits/sec
[  3]  1.0- 2.0 sec  11.2 MBytes  93.8 Mbits/sec
[  3]  2.0- 3.0 sec  11.2 MBytes  93.8 Mbits/sec
[  3]  3.0- 4.0 sec  10.1 MBytes  84.5 Mbits/sec
[  3]  4.0- 5.0 sec  11.2 MBytes  93.8 Mbits/sec
[  3]  5.0- 6.0 sec  11.2 MBytes  93.8 Mbits/sec
[  3]  6.0- 7.0 sec  10.1 MBytes  84.4 Mbits/sec
[  3]  7.0- 8.0 sec  11.2 MBytes  93.8 Mbits/sec
[  3]  8.0- 9.0 sec  11.2 MBytes  93.8 Mbits/sec
[  3]  9.0-10.0 sec  11.3 MBytes  94.6 Mbits/sec
[  3]  0.0-10.1 sec  11.1 MBytes  92.4 Mbits/sec
```

C’est un réseau local de **100Mbits/s**.

8.4.2 IOzone, postmark et tar, ls et rm

IOzone

IOzone est un outil de test d'un système de fichiers. Il génère et mesure une grande variété d'opérations.

Téléchargement

```
# wget http://http.us.debian.org/debian/pool/non-free/i/iozone3/iozone3_287
-2_i386.deb
```

Installation

```
# dpkg –i iozone3_287-2_i386.deb
```

Postmark

Le test effectué par postmark simule l'activité d'une partition de mail avec de nombreuses opérations de création, suppression, concaténation de fichiers de taille variée entre 512 et 10000 octets.

Installation

```
# apt-get install postmark
```

Tar, ls et rm

Ces 3 commandes mesurent respectivement la création, la lecture et la suppression d'un fichier.

8.4.3 Résultats du test

IOzone

(*cf. page 74*)

POSTMARK

Test avec 20000 fichiers et 50000 transactions.

```
# postmark
PostMark v1.51 : 8/14/01
pm>set location /data
pm>set numbers 20000
pm>set transactions 50000
pm>run
```

	Création /sec	Suppression /sec	Transaction /sec
nfs	68	68	105
ocfs2	**352**	**352**	**462**
gfs2	224	224	292

Tableau 3: Résultat du test avec postmark /sec

	Lecture	Ecriture
nfs	224.53 KB/s	423.74 KB/s
ocfs2	**1.13 MB/s**	**2.12MB/s**
gfs2	733.90 KB/s	1.35MB/s

Tableau 4: Résultat du test avec postmark en KB/s ou MB/s

TAR, LS, RM

```
# tar -xzvf linux-2.6.9.tar.gz
# time ls -l linux-2.6.9 > /dev/null
# time rm -rf linux-2.6.9
```

	Création	Lecture	Suppresion
	tar	**ls**	**rm**
nfs	225s	0.130s	53.147s
ocfs2	24s	0.027s	6.780s
gfs2	28s	0.002s	0.497s

Tableau 5: Résultat du test avec tar, ls et rm

8.5 Interprétation des résultats

Le résultat du test effectué avec IOzone ne montre pas vraiment la différence entre les 3 systèmes de fichiers. OCFS2 présente une faiblesse en lecture.

Postmark, simulant le travail d'un serveur Web, montre que OCFS2 est très rapide dans tous les cas, suivit par GFS2.

L'utilisation des commandes Linux montre que OCFS2 est le plus rapide en création de fichiers et GFS2 en lecture et suppression de fichiers. OCFS2 et GFS2 est environ 10 fois plus rapide en création de fichiers.

Vu les résultats des tests effectués, la migration du serveur de stockage NFS vers OCFS2 ou GFS2 est bénéfique en terme de performances.

En plus, du fait que OCFS2 est plus facile à mettre en place et à administrer par rapport à GFS2 sur Debian, le système de fichiers OCFS2 sera utilisé lors de la mise en place du système de stockage distribué (*cf.* **Chapitre 10**, *page 78*).

IOzone

```
# iozone -S 256k -i0 -i1 -i2 -s 1G -r 1M/5M/10M
```

NFS

Taille du fichier de test	Taille du record	Ecriture KB/s	Re-Ecriture KB/s	Lecture KB/s	Re-Lecture KB/s	Lecture aléatoire KB/s	Ecriture aléatoire KB/s
1G	1M	11257	11250	3837017	3625294	3838013	11112
	5M	11681	11120	1780396	1729832	1788823	11978
	10M	11622	11338	1685379	1647075	1688804	11247

Tableau 6: Résultat du test de NFS avec IOzone

OCFS2

Taille du fichier de test	Taille du record	Ecriture KB/s	Re-Ecriture KB/s	Lecture KB/s	Re-Lecture KB/s	Lecture aléatoire KB/s	Ecriture aléatoire KB/s
1G	1M	12168	12119	**41407**	4097615	4234566	12267
	5M	12117	11314	**55641**	1873607	1888677	12425
	10M	12025	12048	**54759**	1776774	1787565	12260

Tableau 7: Résultat du test de OCFS2 avec IOzone

GFS2

Taille du fichier de test	Taille du record	Ecriture KB/s	Re-Ecriture KB/s	Lecture KB/s	Re-Lecture KB/s	Lecture aleatoire KB/s	Ecriture aleatoire KB/s
1G	1M	11785	12239	3879618	3669249	3887588	12891
	5M	12186	11870	1815157	1767028	1793773	13077
	10M	12220	12316	1695414	1653105	1697582	12879

Tableau 8: Résultat du test de GFS2 avec IOzone

Chapitre 9

Gestion du système de fichiers avec LVM

9.1 Initialiser des disques ou des partitions de disques

Avant de pouvoir utiliser un disque ou une partition comme volume physique, il faut l'initialiser :

Pour un disque entier, lancer **pvcreate** sur le disque :

```
# pvcreate /dev/sda
```

Pour les partitions, lancer **pvcreate** sur la partition :

```
# pvcreate /dev/sda1
```

Cela crée un descripteur de groupe de volumes au début de la partition **/dev/sda1**.

Remarques :

L'utilisation d'un disque entier en tant que PV (par rapport à une partition utilisant tout le disque) est déconseillée.

Il est possible d'utiliser un volume RAID pour créer un PV.

9.2 Créer un groupe de volumes

Utiliser le programme **vgcreate** :

```
# vgcreate mon_groupe_de_volumes /dev/sda1
```

9.3 Activer un groupe de volumes

Après un redémarrage ou la commande **vgchange -an**, les VG et LV ne sont plus accessibles. Pour réactiver le groupe de volumes, exécuter :

```
# vgchange -a y mon_groupe_de_volumes
```

9.4 Enlever un groupe de volumes

Assurez-vous qu'aucun volume logique n'est présent dans le groupe de volumes. Désactiver le groupe de volumes :

```
# vgchange -a n mon_groupe_de_volumes
```

Maintenant, vous pouvez supprimer le groupe de volumes :

```
# vgremove mon_groupe_de_volumes
```

9.5 Ajouter un volume physique à un groupe de volumes

Utiliser « **vgextend** » pour ajouter un volume physique déjà initialisé à un groupe de volumes existant.

```
# pvcreate /dev/sdb1
# vgextend mon_groupe_de_volumes /dev/sdb1
```

9.6 Supprimer un volume physique à un groupe de volumes

La commande « **pvdisplay** » permet de s'assurer que le volume physique n'est utilisé par aucun volume logique :

```
# pvdisplay /dev/sda1
--- Physical volume ---
PV Name /dev/sda1
VG Name myvg
PV Size 1.95 GB / NOT usable 4 MB [LVM: 122 KB]
PV# 1
PV Status available
Allocatable yes (but full)
Cur LV 1
PE Size (KByte) 4096
Total PE 499
Free PE 0
Allocated PE 499
PV UUID Sd44tK-9IRw-SrMC-MOkn-76iP-iftz-OVSen7
```

Utiliser ensuite « **vgreduce** » pour enlever le volume physique :

```
# vgreduce mon_groupe_de_volumes /dev/sda1
```

9.7 Créer un volume logique

Pour créer un LV « **testlv** » linéaire de 1500 Mo et son périphérique spécial « **/dev/testvg/testlv** » :

```
# lvcreate -L1500 -ntestlv testvg
```

Pour créer un volume logique de 100 LE avec 2 blocs répartis et une taille de bloc de 4 Ko :

```
# lvcreate -i2 -I4 -l100 -nunautretestlv testvg
```

Pour créer un LV qui utilise tout le VG, utilisez **vgdisplay** pour trouver la valeur de « **Total PE** », puis utilisez-la avec **lvcreate**.

```
# vgdisplay testvg | grep "Total PE"
Total PE 10230

# lvcreate -l 10230 testvg -n monlv
```

Cette dernière commande créera un LV appelé **monlv** qui remplira la totalité du VG **testvg**.

9.8 Supprimer un volume logique

Un volume logique doit être démonté avant d'être supprimé :

```
# umount /dev/monvg/homevol
# lvremove /dev/monvg/homevol
lvremove -- do you really want to remove "/dev/monvg/homevol"? [y/n]: y
lvremove -- doing automatic backup of volume group "monvg"
lvremove -- logical volume "/dev/monvg/homevol" successfully removed
```

9.9 Etendre un volume logique

Pour étendre un volume logique, il suffit de dire à **lvextend** de combien il faut augmenter la taille. Il faut spécifier la quantité d'espace à ajouter ou bien la taille finale du volume logique :

```
# lvextend -L12G /dev/monvg/homevol
lvextend -- extending logical volume "/dev/monvg/homevol" to 12 GB
lvextend -- doing automatic backup of volume group "monvg"
lvextend -- logical volume "/dev/monvg/homevol" successfully extended
```

étend **/dev/monvg/homevol** jusqu'à **12 Go**.

```
# lvextend -L+1G /dev/monvg/homevol
lvextend -- extending logical volume "/dev/monvg/homevol" to 13 GB
lvextend -- doing automatic backup of volume group "monvg"
lvextend -- logical volume "/dev/monvg/homevol" successfully extended
```

ajoute 1 Go à **/dev/monvg/homevol**.

Une fois le volume logique étendu, il est nécessaire d'augmenter la taille du système de fichiers à la taille correspondante. La procédure à suivre dépend du type de système de fichiers utilisé.

Chapitre 10

Trio iSCSI, CLVM et OCFS2

Enfin, ce dernier chapitre explique les étapes à suivre pour la mise en place du système de stockage distribué.

10.1 Préparation

10.1.1 Topologie

La figure suivante montre la topologie utilisée lors du test :

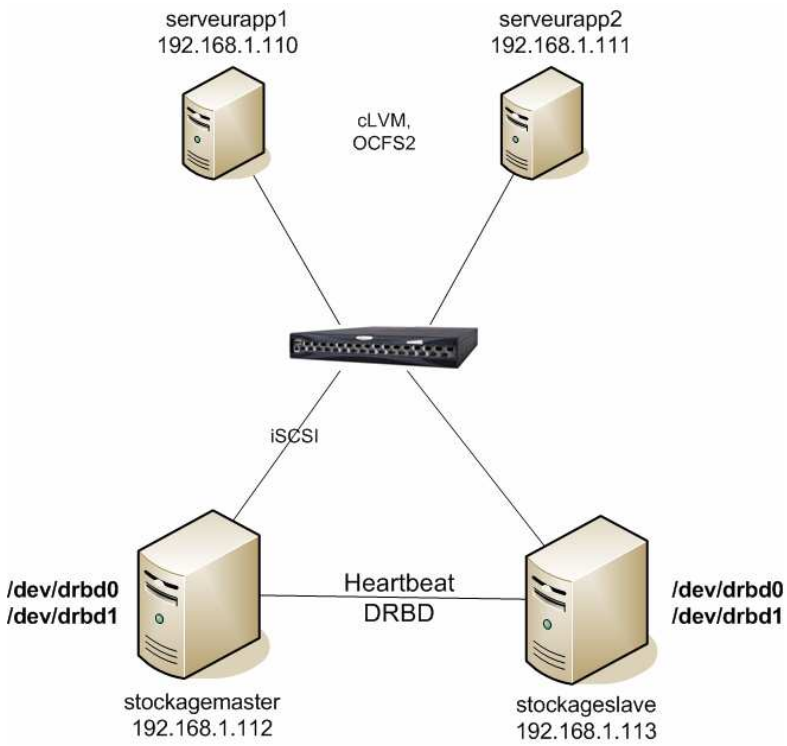

Figure 27: Topologie du test

10.1.2 Installation des machines

Pour réaliser le test, il faut au moins 3 machines. La distribution utilisée est « Debian Lenny ».

10.1.3 Configuration réseau

Il faut configurer les 3 fichiers suivants sur chaque machine:

Le fichier /etc/network/interfaces

```
# This file describes the network interfaces available on your system
# and how to activate them. For more information, see interfaces(5).

# The loopback network interface
auto lo
iface lo inet loopback

auto eth0
iface eth0 inet static
address 192.168.1.11x
netmask 255.255.255.0
broadcast 192.168.1.255
gateway 192.168.1.254
```

Le fichier /etc/hosts

```
127.0.0.1       localhost
192.168.1.110   serveurapp1.malagasy.com        serveurapp1
192.168.1.111   serveurapp2.malagasy.com        serveurapp2
192.168.1.112   stockagemaster.malagasy.com     stockagemaster
192.168.1.113   stockageslave.malagasy.om       stockageslave
```

Le fichier /etc/resolv.conf

```
domain malagasy.com
search malagasy.com blueline.mg
nameserver 41.204.103.205
nameserver 41.204.104.45
```

10.2 Importer le disque avec iSCSI

10.2.1 Côté iSCSI Target

Pour l'installation et la configuration, *cf.* **iSCSI Target**, *page 67*. Le disque à exporter est **/dev/drbd0**.

10.2.2 Côté iSCSI Initiator

Pour l'installation et la configuration, *cf.* **iSCSI Initiator**, *page 68*.

10.2.3 Partitionner le nouveau disque

Découvrir le nouveau disque:

```
# fdisk –l
...
```

Partitionner:

```
# fdisk /dev/sdb
# fdisk –l /dev/sdb
```

10.3 Installation et configuration de CLVM

Par défaut, CLVM utilise CMAN (Cluster MANager) comme gestionnaire de cluster. L'administration de CMAN sous Debian est difficile. Pour ce faire, il faut compiler CLVM pour qu'il utilise OpenAIS par défaut.

10.3.1 OpenAIS, une alternative à CMAN avec CLVM

Installer toutes les dépendances nécessaires pour compiler CLVM :

```
# apt-get install dpkg-dev
# apt-get build-dep clvm
# apt-get install libopenais-dev
```

Télécharger les fichiers sources:

```
# cd /usr/src
/usr/src # apt-get source clvm
/usr/src # cd lvm2-2.02.39
```

Modifier 3 fichiers sources :

- Le premier fichier est **debian/clvm.init**. Il faut supprimer toutes les références à **cman** et **cluster.conf**.

- Le second est **debian/control**. Modifier les dépendances (supprimer le numéro de version de lvm2, et remplacer cman par openais) et modifier les commentaires.

- Le dernier fichier est **debian/rules**. Remplacer **cman** par **openais** dans l'option de configurations, et ajouter le chemin de la librairie d'openais.

Voici un extrait du fichier **debian/rules** :

```
$(STAMPS_DIR)/setup-deb: SOURCE_DIR = $(BUILD_DIR)/source
$(STAMPS_DIR)/setup-deb: DIR = $(BUILD_DIR)/build-deb
$(STAMPS_DIR)/setup-deb: $(STAMPS_DIR)/source
      rm -rf $(DIR)
      cp -al $(SOURCE_DIR) $(DIR)
      cd $(DIR); \
      /configure CFLAGS="$(CFLAGS)" \
      LDFLAGS="-L/usr/lib/openais" \
            $(CONFIGURE_FLAGS) \
            --with-optimisation="" \
            --with-clvmd=openais \
```

Cf. **Annexe A**, *page 91* pour plus de détails sur les 2 autres fichiers.

Metter à jour la version du paquet clvm :

```
/usr/src/lm2-2.02.39# apt-get install devscripts
/usr/src/lm2-2.02.39# dch -i
```

Compiler le paquet maintenant :

```
/usr/src/lm2-2.02.39# dpkg-buildpackage -rfakeroot -uc -b
```

Après la compilation, le nouveau paquet **clvm*.deb** se trouve dans le répertoire **/usr/src**.

Installer openais et ajouter un utilisateur :

```
# apt-get install openais
# mkdir -p /etc/ais
# adduser --no-create-home --disabled-password --disabled-login --gecos
openAIS ais
```

Le fichier de configuration d'openais est **/etc/ais/openais.conf.**

```
# nano /etc/ais/openais.conf
totem {
      version: 2
      secauth: off
      threads: 0
      interface {
            ringnumber: 0
            bindnetaddr: 192.168.1.0
            mcastaddr: 226.94.1.1
            mcastport: 5405
      }
}
```

Créer le script **openais** (*cf.* **Annexe B**, *page 95*)

Lancer le script openais au démarrage :

```
# update-rc.d openais start 62 S . start 50 0 6 .
```

Installer clvm et activer LVM en cluster :

```
/usr/src # cd lvm2*.deb
/usr/src # dpkg –i clvm*.deb
/usr/src # nano /etc/lvm/lvm.conf
locking_type = 3
```

Démarrer/Arrêter les services :

```
# /etc/init.d/clvm stop
# /etc/init.d/lvm2 stop
# /etc/init.d/openais start
# /etc/init.d/clvm start
# /etc/init.d/lvm2 start
```

10.3.2 Créer un PV, VG et LV

Créer un PV

```
# pvcreate /dev/sdb1
```

Créer un VG

```
# vgcreate vg_storage1 /dev/sdb1
```

Créer un LV

```
# lvcreate –L 100%FREE –n lv_mail vg_storage1
```

Configurer le fichier **/etc/ocfs2/cluster.conf** :

Formater :

```
# mkfs.ocfs2 –L mail /dev/vg_storage1/lv_mail
```

Vérifier:

```
# pvscan
# vgscan
# lvdisplay
```

10.4 Ajouter un autre nœud

Pour ajouter un autre nœud au cluster, c'est-à-dire le **serveurapp2**, il faut refaire les étapes :

- **OpenAIS, une alternative à CMAN avec CLVM**, *cf. page 80* et
- **Importer le disque avec iSCSI**, *cf. page 79*.

```
# fdisk -l
# pvscan
# vgscan
# lvdisplay
```

Le LV **`/dev/vg_storage1/lv_mail`** est visible sur tous les nœuds du cluster.

Remarque :

Le disque **/dev/sdb1** sur le serveur de stockage **stockagemaster**, importé de nouveau sur le serveur « **serveurapp2** » n'a plus besoin d'être partitionné ni formaté.

10.5 Ajouter un autre disque

Voici les étapes à suivre pour ajouter un autre disque:

- Préparer le disque sur un iSCSI Target existant et exporter le,
- Importer le disque sur l'iSCSI Initiator,
- Créer un PV avec le nouveau disque importé,
- Ajouter le nouveau PV à un VG
- Etendre le LV.

Serveur **stockagemaster** :

Le disque à exporter est **/dev/drbd1** (*cf.* **iSCSI Target**, *page 67*). Ajouter le dans le fichier **/etc/ietd.conf**.

Serveur **serveurapp1** :

Exporter le nouveau disque (*cf.* **iSCSI Initiator**, *page 68*)

Découvrir le nouveau disque et partitionner le:

```
# fdisk –l
# fdisk /dev/sdc
# fdisk –l /dev/sdc
```

Créer un PV, ajouter le dans vg_storage, étend la taille du LV et du système de fichiers :

```
# pvcreate /dev/sdc1

# vgextend vg_storage1 /dev/sdc1

# lvextend –l 100%FREE /dev/vg_storage1/lv_mail

# tunefs.ocfs2 –S /dev/vg_storage1/lv_mail
```

Remarques:

Le changement au niveau du LV est immédiatement visible au sein du cluster.

Les étapes à suivre lors de l'ajout d'un autre serveur de stockage sont les mêmes que lors de l'ajout d'un autre disque.

CONCLUSION

Ainsi, j'ai effectué mon stage au sein de la société BLUELINE. Lors de ce stage de 5 mois, j'ai pu mettre en pratique mes connaissances théoriques acquises durant ma formation. De plus, je me suis confronté aux difficultés réelles du monde du travail.

Aujourd'hui, dans la plupart des infrastructures informatiques, l'explosion des volumes de données pose un réel problème, dont la solution doit être étudiée de façon globale.

Les solutions NAS ou SAN ouvrent des nouvelles perspectives et font du réseau un acteur prépondérant dans le service des données. Mais stocker un volume important n'est pas la seule question. Le réel enjeu se situe au niveau des services associés à ces données : les temps de réponse, la disponibilité et la sécurité.

Le but de ce stage était d'étudier les différentes possibilités et de choisir la meilleure solution. Apres une étude approfondie et en tenant compte de plusieurs critères, la solution SAN est retenue. Cette solution est évolutive, souple, performante et non coûteuse.

Pour la réalisation pratique, il est important de noter l'utilisation de la distribution Debian GNU Linux, un système d'exploitation composé uniquement de logiciels libres.

Enfin, il est important de noter que contrairement aux distributions "RedHat" et "Suse", qui possèdent une version dédiée pour la haute disponibilité, ce travail contribue à la mise en place d'un tel système sous Debian.

BIBLIOGRAPHIE

[1] : Abdelwaheb DIDI, Gilles SCALA et Michael DIOT; Les solutions de stockage en réseau, le SAN contre le NAS,
2003.

[2] : David Komar, Guillaume Le Cam, Mathieu Mancel; Les solutions de stockage NAS et SAN, 11 Janvier 2006.

[3] : Cyrille DUFRESNES, SAN et NAS
18 Août 2008

[4] : Thomas Briet et Cyril Muhlenbach, Stockage DAS/SAN/NAS/iSCSI
2000

[5] : Pierre Olivier BOURGEOIS et Alexis MARCOU, Storage Area Network
26 Juillet 2004

[6] : A. J. Lewis, Guide pratique de LVM
28 Janvier 2007

[7] : Philippe Latu, Introduction au réseau de stockage iSCSI
16 Mai 2010.

[8] : Novell SUSE Linux Enterprise Server, Storage Administration Guide
23 Février 2010

WEBOGRAPHIE

iSCSI

[9]: http://blogpmenier.dynalias.net/docext/ocfs2/

GNBD

[10]: http:/docs.redhat.com/docs/fr-FR/Red_Hat_Enterprise_Linux/5/html/Cluster_Suite_Overview/s1-gnbd-overview-CSO.html
[11]: http://www.wains.be/index.php/2009/06/30/gnbd-on-debian-installation-howto/
[12]: http://blogpmenier.dynalias.net/docext/gnbd/

LVM

[13]: http://www.admin-debian.com/les-systemes-de-fichiers-linux/lvm-2-logical-volume-management/

CLVM

[14]: http://www.pixelchaos.net/2009/04/23/openais-an-alternative-to-clvm-with-cman/
[15]: http://h2o.glou.fr/post/2009/04/20/clvm-openais-on-Debian/Lenny
[16]: http://docs.redhat.com/en-US/Red_Hat_Enterprise_Linux/5/html/Cluster_Logical_Volume_Manager/LVM_Cluster_Overview.html

NFS

[17]: http://www.howtoforge.com/nfs-server-and-client-debian-etch
[18]: http://www.lininfo.org/spip.php?article13

OCFS2

[19]: http://oss.oracle.com/projects/ocfs2/dist/documentation/v1.2/ocfs2_user_guide.pdf

GFS2

[20]: http://docs.redhat.com/docs/fr-FR/Red_Hat_Enterprise_Linux/5/html/Global_File_System_2/index.html

Bonnie++

[21]: http://jfenal.free.fr/Traduc/LG68F/lg68f-fr/lg68f-7.fr.html

Postmark

[22]: http://www.linux-france.org/article/sys/ext3fs/Benchmarks/PostMark.txt
[23]: http://www.linux-france.org/article/sys/ext3fs/Benchmarks/petit-test.txt

TABLE DES MATIERES

ANNEXES

Annexe A

Le fichier **debian/clvm.init**

```
#!/bin/sh
#
### BEGIN INIT INFO
# Provides:          lvm cluster locking daemon
# Required-Start:    lvm2
# Required-Stop:     lvm2
# Default-Start:     2 3 4 5
# Default-Stop:      0 1 6
# Short-Description: start and stop the lvm cluster locking daemon
### END INIT INFO
#
# Author: Frederik Schüler <fs@debian.org>
# based on the old clvm init script from etch
# and the clvmd init script from RHEL5

PATH=/sbin:/usr/sbin:/bin:/usr/bin
DESC="Cluster LVM Daemon"
NAME=clvm
DAEMON=/sbin/clvmd
SCRIPTNAME=/etc/init.d/clvm

[ -x $DAEMON ] || exit 0

. /lib/init/vars.sh

. /lib/lsb/init-functions

CLVMDTIMEOUT=20

# Read configuration variable file if it is present
[ -r /etc/default/$NAME ] && . /etc/default/$NAME

DAEMON_OPTS="-T$CLVMDTIMEOUT"

do_start()
{
      start-stop-daemon --start --quiet --exec $DAEMON -- $DAEMON_OPTS ||
status="$?"
      # flush cache
      vgscan > /dev/null 2>&1
      return $status
}

do_activate()
{
      if [ -n "$LVM_VGS" ] ; then
            log_action_msg  "Activating VGs $LVM_VGS"
            vgchange -ayl $LVM_VGS || return $?
      else
            log_action_msg  "Activating all VGs"
            vgchange -ayl || return $?
      fi
```

```
}

do_deactivate()
{
      if [ -n "$LVM_VGS" ] ; then
            vgs="$LVM_VGS"
      else
            # Hack to only deactivate clustered volumes
            vgs=$(vgdisplay -C -o vg_name,vg_attr --noheadings 2> /dev/null
| awk '($2 ~ /.....c/) {print $1}')
      fi

      [ "$vgs" ] || return 0

      vgchange -anl $vgs || return $?
}

do_stop()
{
      start-stop-daemon --stop --quiet --name clvmd
      status=$?
      return $status
}

case "$1" in
      start)
            # start the daemon...
            log_daemon_msg "Starting $DESC" "$NAME"
            do_start
            status=$?
            case "$status" in
                  0) log_end_msg 0 ;;
                  1) log_action_msg " already running" ; log_end_msg 0 ;;
                  *) log_end_msg 1 ;;
            esac
            # and activate clustered volume groups
            do_activate
            status=$?
            exit $status
      ;;
      stop)
            # deactivate volumes...
            log_daemon_msg "Deactivating VG $vg:"
            do_deactivate
            status=$?
            case "$status" in
                  0) log_end_msg 0 ;;
                  1) log_end_msg 0 ;;
                  *) log_end_msg 1 ;;
            esac
            # and stop the daemon
            log_daemon_msg "Stopping $DESC" "$NAME"
            do_stop
            status=$?
            case "$status" in
                  0) log_end_msg 0 ; exit 0 ;;
                  1) log_end_msg 0 ; exit 0 ;;
                  *) log_end_msg 1 ; exit $status ;;
            esac
      ;;
      restart|force-reload)
```

```
            $0 stop
            sleep 1
            $0 start
      ;;
      status)
            pid=$( pidof $DAEMON )
            if [ -n "$pid" ] ; then
                  log_action_msg "$DESC is running"
            else
                  log_action_msg "$DESC is not running"
            fi
            exit 0
      ;;
      *)
            echo "Usage: $SCRIPTNAME {start|stop|restart|force-
reload|status}" >&2
            exit 1
      ;;
esac

exit 0
```

Le fichier **debian/control**

```
Source: lvm2
Section: admin
Priority: optional
Maintainer: Debian LVM Team <pkg-lvm-maintainers@lists.alioth.debian.org>
Uploaders: Bastian Blank <waldi@debian.org>
Build-Depends: debhelper (>> 4.2), automake, libcman-dev (>> 2),
libdevmapper-dev (>> 2:1.02.27), libdlm-dev (>> 2), libreadline5-dev, quilt
Standards-Version: 3.7.3

Package: lvm2
Architecture: any
Depends: ${shlibs:Depends}, lsb-base
Conflicts: lvm-common
Replaces: lvm-common
Recommends: dmsetup
Description: The Linux Logical Volume Manager
 This is LVM2, the rewrite of The Linux Logical Volume Manager.  LVM
 supports enterprise level volume management of disk and disk subsystems
 by grouping arbitrary disks into volume groups. The total capacity of
 volume groups can be allocated to logical volumes, which are accessed as
 regular block devices.

Package: lvm2-udeb
XC-Package-Type: udeb
Section: debian-installer
Architecture: any
Depends: ${shlibs:Depends}
Description: The Linux Logical Volume Manager
 This is a udeb, or a microdeb, for the debian-installer.
 .
 This is LVM2, the rewrite of The Linux Logical Volume Manager.  LVM
 supports enterprise level volume management of disk and disk subsystems
 by grouping arbitrary disks into volume groups. The total capacity of
 volume groups can be allocated to logical volumes, which are accessed as
 regular block devices.
```

```
Package: clvm
Section: admin
Priority: extra
Architecture: any
Depends: ${shlibs:Depends}, lvm2, lsb-base, openais
Description: Cluster LVM Daemon for lvm2
 This package provides the clustering interface for lvm2. It is a modified
 version that uses openais instead of Red Hat's "cman" cluster
infrastructure. It allows logical volumes to be created on shared storage
devices (eg Fibre Channel, or iSCSI).
```

Annexe B

Le script **openais**

```
#! /bin/sh
#
# openais init.d script
#
# update-rc.d openais start 62 S . start 50 0 6 .
#

PATH=/usr/local/sbin:/usr/local/bin:/sbin:/bin:/usr/sbin:/usr/bin
DAEMON=/usr/sbin/aisexec
NAME=openais
DESC=openais
FLAGS=

test -f $DAEMON || exit 0

set -e

JOIN_TIMEOUT=15

# Read configuration variable file if it is present
[ -r /etc/default/$NAME ] && . /etc/default/$NAME

case "$1" in
  start)
      echo -n "Starting $DESC: "
      start-stop-daemon --start --quiet -o --exec $DAEMON -- $FLAGS
      time=0
      while [ "$JOIN_TIMEOUT" -eq 0 ] || [ "$time" -lt "$JOIN_TIMEOUT" ] ;
do
            sleep 1
            if openais-cfgtool -s &>/dev/null ; then
                  echo "$NAME."
                  exit 0
            else
                  echo -n " . "
                  time=$(($time + 1))
            fi

      done
      echo "FAILED"
      exit 1
      ;;
  stop)
      echo -n "Stopping $DESC: "
      start-stop-daemon --stop --quiet -o --exec $DAEMON
      echo "$NAME."
      ;;
  reload|force-reload)
       echo "Reloading $DESC configuration files."
       start-stop-daemon --stop --signal 1 --quiet -o --exec $DAEMON
      ;;
  restart)
      echo -n "Restarting $DESC: "
      start-stop-daemon --stop --quiet -o --exec $DAEMON
      sleep 1
      start-stop-daemon --start --quiet -o --exec $DAEMON -- $FLAGS
```

```
      echo "$NAME."
      ;;
  *)
      N=/etc/init.d/$NAME
      echo "Usage: $N {start|stop|restart|reload|force-reload}" >&2
      exit 1
      ;;
esac

exit 0
```

GLOSSAIRE

DRBD

DRBD (Distributed Replicated Block Device) est un mécanisme de réplication de données localisées sur deux serveurs distincts par voie réseau. Pour simplifier, il s'apparente à du RAID-1 sur IP. Quand une écriture a lieu sur le disque du serveur maître, l'écriture est simultanément réalisée sur le serveur esclave. La synchronisation est faite au niveau de la partition.

Cluster

Ordinateurs en grappe qui partagent le travail et/ou peuvent prendre le relais les uns des autres. Une des ces machines constitue un nœud du cluster.

Commutateur

Un commutateur réseau (ou switch, de l'anglais) est un équipement qui relie plusieurs segments (câbles ou fibres) dans un réseau informatique. Il s'agit le plus souvent d'un boîtier disposant de plusieurs ports Ethernet (entre 4 et plusieurs centaines).

Conteneur

Le conteneur est une « virtualisation au niveau du système d'exploitation ». Le but est d'isoler des environnements de l'espace utilisateur entre eux mais en les faisant fonctionner sur le même noyau. Ces systèmes de conteneur permettent alors d'utiliser différents ensembles d'applications dans un même contexte (on peut alors envisager plusieurs distributions Linux avec un unique noyau).

Fibre Channel

Technologie de transmission de données entre ordinateurs à haut débit (jusqu'à 1 Go/s). Elle est particulièrement employée pour connecter des unités de stockages aux serveurs. On peut utiliser le Fibre Channel sur plusieurs supports physiques : fibre optique (pour les grandes distances (10 km)), câble coaxial, paire torsadée.

GFS2

Global File System ou Système de fichiers global est un système de fichiers partagé conçu pour une grappe d'ordinateurs Linux ou IRIX. GFS et GFS2 sont différents des systèmes de fichiers distribués comme AFS, Coda, ou InterMezzo parce qu'ils permettent à tous les nœuds

un accès direct concurrent au même périphérique de stockage de type bloc. De plus, GFS et GFS2 peuvent également être utilisés comme un système de fichiers local.

Heartbeat

Application permettant à un ordinateur de prendre le pouls (Heartbeat) d'autres machines. Si l'une d'entre elles ne répond pas à un message envoyé, elle est considérée comme défaillante ; une mesure de secours est alors prise.

LUN

Le Logical Unit Number est une unité logique qui convertit un espace brut d'espace disque en unité de stockage logique que le système d'exploitation d'un serveur peut accéder et utiliser.

NAS

Un stockage en réseau NAS, ou un NAS (de l'anglais Network Attached Storage), est un serveur de fichiers autonome, reliée à un réseau dont la principale fonction est le stockage de données en un volume centralisé pour des clients réseau hétérogènes.

NFS

NFS (Network File System) est un protocole permettant de monter des disques en réseau. Ce protocole basé sur le principe client/serveur, a été développé par Sun Microsystems en 1984. Il peut servir à l'échange de données entre des systèmes Linux, Mac ou Windows. L'un de ses avantages est qu'il gère les permissions sur les fichiers.

OCFS2

OCFS2 a été développé par Oracle Corporation, au début seulement pour les bases de données Oracle mais dans sa deuxième version, il permet d'être utilisé pour de différentes applications. Pour faire la gestion de redondance et de distribution de l'information, il utilise de battements de cœur pour savoir quels sont les nœuds présents dans le cluster.

OpenVZ

OpenVZ est une méthode de virtualisation de serveur développée à partir de Linux. OpenVZ crée des serveurs privés virtuels (VPS) sécurisés et isolés ou des environnements virtuels sur un unique serveur physique permettant une meilleure utilisation du serveur et s'assurant qu'il n'y ait aucun conflit entre les applications.

RAID

Moyen de stocker des données à différents endroits sur plusieurs disques durs. L'ensemble apparaît au système sous la forme d'un seul périphérique de stockage. Selon la configuration (RAID-1, RAID-5), la vitesse de lecture, la tolérance aux pannes, la correction d'erreurs peut être améliorée.

Routeur

Un routeur est un élément intermédiaire dans un réseau informatique assurant le routage des paquets. Son rôle est de faire transiter des paquets d'une interface réseau vers une autre, selon un ensemble de règles formant la table de routage. C'est un équipement de couche 3 du modèle OSI.

SAN

Un réseau de stockage, ou SAN (de l'anglais Storage Area Network), est un réseau spécialisé permettant de mutualiser des ressources de stockage.

SCSI

Le SCSI est un Standard décrivant une interface parallèle permettant aux ordinateurs de communiquer avec leurs périphériques (de stockage notamment).

SPOF (Single Point Of Failure)

C'est un point d'un système informatique dont le reste du système est dépendant. Autrement dit, une panne au niveau de ce SPOF entraîne un arrêt complet du système informatique. La présence de SPOF dans un système informatique augmente la probabilité d'apparition d'un déni de service et entraîne un risque sur la qualité de service.

Serveur virtuel

Le serveur virtuel est une machine physique que l'on découpe en serveur autonome complet. Toute la souplesse du serveur dédié avec des capacités plus restreinte, mais plus de sécurité et un coût plus compétitif.

Snapshot (clichés)

Les *snapshots* sont des volumes logiques permettant d'effectuer une sauvegarde cohérente d'un autre volume logique du même groupe de volumes. La création d'un *snapshot* consiste à prendre une « photo », un instantané du volume logique cible (ce qui est quasi-immédiat) et

on commence alors à enregistrer les modifications apportées au volume logique cible.

Système de fichiers distribué

Un système de fichiers distribué ou système de fichiers en réseau est un système de fichier qui permet le partage de données à plusieurs clients au travers du réseau. Contrairement à un système de fichier local, le client n'a pas accès au système de stockage sous-jacent, et interagit avec le système de fichier via un protocole adéquat.

Virtualisation

Le principe de la virtualisation consiste à faire fonctionner plusieurs « serveurs » sur une seule machine physique. On entend par « serveur », l'ensemble « Système d'exploitation » et « Applications ».

Virtualisation des serveurs

La virtualisation des serveurs permet de partager un même serveur entre plusieurs applications, systèmes d'exploitations ou utilisateurs. Elle permet également le transfert d'applications d'un serveur à l'autre.

Virtualisation du stockage

Appelé aussi abstraction du stockage, il sépare la présentation logique et la réalité physique des ressources de stockage et permet ainsi de gérer le stockage indépendamment du matériel de stockage physique sous-jacent.

Zoning

Le zoning est une méthode d'organisation des périphériques connectés en Fibre Channel en groupes logiques à partir de la configuration physique. Elle permet la limitation et le contrôle du trafic entre l'adaptateur FC installé sur le serveur et l'unité de stockage attachés. Le zoning peut être de type matériel ou logiciel.

RESUME

Actuellement, l'explosion des volumes de données à gérer est une réalité au sein des entreprises y compris la société BLUELINE où j'ai effectue mon stage.

Ce stage a pour but de mettre en place un système de stockage distribué pour résoudre ce problème. Le réseau de stockage SAN a été choisit parmi toutes les solutions possibles. SAN offre une gestion simplifiée de stockage, l'évolutivité, la flexibilité, la disponibilité et l'amélioration de l'accès aux données et de sauvegarde. De plus, cette solution est totalement libre, puisqu'elle utilise des logiciels libres avec la distribution Debian.

Après l'étude de l'existant, qui a permis de mieux comprendre les avantages et les inconvénients du système existant, une étude approfondie a été faite. Cette étude aboutit à l'utilisation du trio iSCSI, CLVM et OCFS2. iSCSI pour exporter des disques en réseau, CLVM pour faciliter la gestion de l'espace de stockage et OCFS2 gère l'écriture et la lecture sur un disque dans un environnement en cluster.

Durant le stage, la réalisation pratique a été effectuée sur des serveurs de test avant de le mettre en production. Les résultats ont été concluant.

Mots clés : stockage distribué**,** iscsi, gnbd, lvm2, clvm, nfs, ocfs2, gfs2, bonnie++, postmark

ABSTRACT

The explosion of data created by the businesses of today is making storage a strategic investment priority for companies of all sizes including BLUELINE company where I did my training.

This training aims to set up a shared storage infrastructure to solve this problem. The SAN was chosen among all possible solutions. SANs offer simplified storage management, scalability, flexibility, availability, and improved data access and backup. Moreover, this solution is totally free; it uses open software under the Debian distribution.

After studying the existing, which led to a better understanding of the advantages and disadvantages of the existing system, a detailed study was made. This study leads to the use of iSCSI, CLVM and OCFS2. iSCSI exports disks on the network, CLVM manages the shared storage and OCFS2 allows all nodes in the cluster to access the filesystem concurrently.

During the training, the practical test has been done on test servers before putting it into production. The results were conclusive.

Key words: shared storage, iscsi, gnbd, lvm2, clvm, nfs, ocfs2, gfs2, bonnie++, postmark