Mise en oeuvre d'un nouveau système de stockage SAN

Jean-Edouard Simonin

▶ **To cite this version:**

Jean-Edouard Simonin. Mise en oeuvre d'un nouveau système de stockage SAN. Sciences de l'ingénieur [physics]. 2012. hal-02094481

HAL Id: hal-02094481

https://hal.univ-lorraine.fr/hal-02094481

Submitted on 9 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-memoires-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

Université de Lorraine
Faculté des Sciences et Technologies

Master Ingénierie de Systèmes Complexes
Spécialité Réseaux, Signaux, Images
Année universitaire 2011-2012

Mise en œuvre d'un nouveau système de stockage SAN

Mémoire présenté par Jean-Edouard SIMONIN
Soutenu le vendredi 07 septembre 2012

Stage effectué à la Direction du Système d'Information du CHU de Nancy
92 Avenue du Maréchal De Lattre De Tassigny
54000 NANCY

Tuteur industriel : Francis DAUL
Tuteur universitaire : Thierry BASTOGNE

Remerciements

Tout d'abord, je tiens à remercier tout particulièrement M Francis DAUL, Responsable d'Exploitation, qui a répondu favorablement à ma candidature et qui a su me faire confiance malgré la criticité des données que j'avais à manipuler. Sa grande disponibilité m'a permis de me sentir immédiatement intégré et à l'aise.

Je tiens à exprimer toute ma gratitude à M Alban COUJARD, Ingénieur d'Intégration, pour la pédagogie et la patience dont il a su faire preuve à mon égard et au contact duquel j'ai beaucoup appris.

Je veux aussi remercier M Bruno FRANCOIS, Analyste d'Exploitation, pour sa disponibilité et ses conseils techniques.

Je remercie aussi l'ensemble de l'équipe de la DSI et plus particulièrement du secteur Exploitation pour leur accueil sympathique et leur coopération professionnelle durant ces cinq mois.

Je tiens enfin à témoigner toute ma reconnaissance à mon tuteur universitaire, M Thierry BASTOGNE, pour ses conseils avisés lors de la visite de stage.

Introduction

Dans le cadre de l'UE MAIC4UP, j'ai effectué un stage de cinq mois au sein de la Direction du Système d'Information du CHU de Nancy. Ce choix de stage fait suite à la formation de découverte proposée par M RABITA dans l'UE S2AB. Elle a été pour moi une révélation et ce stage était l'occasion de conforter mon choix d'orientation qu'est le domaine du stockage. Il m'a semblé donc pertinent de candidater au CHU de Nancy, plus gros employeur de Meurthe-Et-Moselle, et qui dispose d'un système d'information en conséquence.

J'ai ainsi intégré le secteur Exploitation et ai participé activement à la mise en œuvre d'un nouveau système de stockage SAN. Sous la tutelle de M DAUL et en collaboration avec M COUJARD, j'ai pris part à un projet de grande ampleur consistant à mettre en place les nouvelles baies de stockage sur les deux sites que compte le CHU. Il a fallu ensuite migrer les données des anciennes vers les nouvelles baies en assurant tout au long de ce processus une continuité de service.

Sommaire

Remerciements **2**

Introduction **3**

1. Présentation **5**

1.1 Le CHU de Nancy 5

1.2 La DSI 7

2. L'Exploitation **8**

2.1 Démarche qualité ITIL 8

2.2 Le centre de services 8

2.3 Intégration technique des applications 9

2.3.1 Automatisation et industrialisation de la production 9

2.3.2 Urbanisation 9

2.4 Télémaintenance 10

2.5 Salles informatiques 10

2.6 Architecture de stockage 11

2.6.1 SAN 11

2.6.2 NAS 15

2.7 Architecture de Sauvegarde 15

3. Le projet **16**

3.1 Présentation 16

3.1.1 Sujet de stage 16

3.1.2 Contexte 16

3.1.3 Nouvelle configuration matérielle 17

3.1.4 Etapes de la migration 18

3.2 Gestion de projet 19

3.2.1 Les besoins client 19

3.2.2 Les solutions choisies 22

3.3 Travaux réalisés 22

3.3.1 Outil d'aide à la création de LUNs 22

3.3.2 Métrologie 25

Conclusion **28**

Bibliographie **28**

1. Présentation

1.1 Le CHU[1] de Nancy

Le CHU de Nancy s'appuie sur 5 sites répartis sur 2 pôles, Nancy et Brabois :

- Hôpital Central : Mis en service en 1883, et complété par le nouveau bâtiment neurologique en 1999, il héberge le dispositif d'urgence de l'agglomération nancéienne : Service Accueil des Urgences, centre 15, SAMU, centre antipoison. Il accueille principalement dans ses 444 lits et places :
  - des services de médecine interne, de neurologie et de réanimation médicale.
  - des services de chirurgie pour les spécialités d'orthopédie et de traumatologie, de neurochirurgie, d'ophtalmologie, d'ORL et maxilo-faciale, avec une réanimation chirurgicale.
  - un important plateau technique d'imagerie et de laboratoires : IRM, scanner...

Hôpital Central

Bâtiment de Neuroscience

- Hôpital Saint Julien : Structure d'hébergement de 94 lits pour personnes âgées, de sa création en 1900 à nos jours, l'hôpital Saint-Julien offre le calme de ses jardins en plein centre-ville de Nancy pour les soins de longue durée.

Hôpital Saint Julien

- Centre Saint Stanislas : Mis en service fin 1993, le nouveau centre de long séjour s'adresse aux personnes âgées en complément de l'Hôpital Saint Julien.

[1] Centre Hospitalier Universitaire

Centre Saint Stanislas

- Hôpital d'adultes de Brabois : Inauguré en 1973, il constitue un ensemble hospitalo-universitaire de premier plan avec la proximité de la faculté de médecine et le Centre Régional de Lutte Contre le Cancer avec un potentiel de 945 lits et places.

Hôpital d'adultes de Brabois

Institut lorrain du cœur et des vaisseaux

- Hôpital d'enfants de Brabois : L'hôpital d'enfants a ouvert ses portes en 1982 offrant les meilleures conditions d'accueil aux hospitalisés et à leurs familles. Il comprend dans ses 227 lits et places :
    - des services de médecine pédiatriques spécialisés (cancérologie, génétique chimique, pédo-psychiatrie…).
    - des services de chirurgie spécialisés en orthopédie et en viscérale.
    - un plateau technique spécifique.
    - le Pôle d'Accueil Spécialisé d'Urgence pour la spécificité infantile.

Hôpital d'enfants de Brabois

Ces cinq établissements accomplissent la mission première d'un CHU : les soins. Ils permettent chaque année plus de 430000 consultations externes et 66000 accueils aux urgences. Les 1700 lits, quant à eux, autorisent 70000 patients.

A cette mission s'en ajoutent deux autres : l'enseignement et la recherche.

Pour la première, ce sont 10 établissements qui ne forment pas moins de 6000 professionnels tous les ans (aides-soignants, ambulanciers, infirmiers,…). La recherche et

l'innovation, pour leur part, soutiennent plus de 2500 publications sur 850 projets concrétisés par 12 brevets. Cela fait de Nancy le 10ème CHU de France dans ce domaine.

Au total, le Centre Hospitalier compte plus de 9000 salariés, ce qui en fait le plus gros employeur de Meurthe-Et-Moselle. Côté logistique, ce sont par exemple 6000 repas par jour servis ou près de 3500 tonnes de linge traitées par an. Ainsi, en 2010, ses dépenses se sont élevées à 636 M€ (soit 1.74 M€ par jour) pour 620.6 M€ de recettes.

1.2 La DSI[2]

La DSI se situe à l'Hôpital Marin. Annexe de l'Hôpital Central, l'Hôpital Marin n'abrite plus aucun service médical depuis le départ en 1973 du service de Chirurgie pour Brabois, mais uniquement des services administratifs. La DSI compte 70 informaticiens répartis en deux départements.

DÉPARTEMENT FONCTIONNEL	
GRH-PAIE-GTA Dominique THOMAS	GEF- CIRCUIT DU MÉDICAMENT -GMAO- INFOCENTRE Évelyne COURTIAL
MÉDICO-TECHNIQUE Sébastien TRIQUENOT	GAM-GML-FACTURATION Joël PARSAD
INTEROPÉRABILITÉ Didier PETITCOLAS	DOSSIER PATIENT - VALORISATION Murielle HAZEMANN

DÉPARTEMENT PRODUCTION	
SYSTÈMES Béatrice BÉRARD	RÉSEAU-TÉLÉPHONIE Éric GUESNEY
EXPLOITATION Francis DAUL	MICRO-INFORMATQIUE Gérard PERRAUD
PACS Claude LEFONDEUR	PSSI Jean-Stanislas TYZO

Organigramme de la Direction du Système d'Information

Le département fonctionnel assure l'intégration fonctionnelle des applications informatiques : rédaction des cahiers des charges en concertation avec l'utilisateur hospitalier, adaptation des logiciels au contexte pour une meilleure intégration et support fonctionnel niveau 2.

Le département production assure quant à lui l'intégration technique des applications informatiques et l'industrialisation de leur exploitation. Il compte en ses rangs plusieurs secteurs :

- Systèmes : gestion de l'architecture des serveurs, les clients légers, la messagerie et l'annuaire d'entreprise
- Réseau/téléphonie : gestion de la téléphonie et le réseau VDI[3]
- Micro-informatique : gestion du poste de travail (le PC, l'imprimante,…)
- PACS[4] : informatisation de l'imagerie
- PSSI : gestion de la Politique de Sécurité du Système d'Information

Enfin, il compte le secteur Exploitation que j'ai intégré durant ce stage.

[2] Direction du Système d'Information
[3] Voix, Données, Image
[4] Picture Archiving and Communication Systems ou dossier patient de radiologie

2. L'Exploitation

Le secteur Exploitation assure le bon fonctionnement et l'optimisation de la production informatique du CHU de Nancy. Pour ce faire, il développe l'automatisation qui repose sur l'industrialisation de la production.

Il veille au respect des calendriers de traitements, à la disponibilité des applications et à la qualité de service.

Les activités gérées par le secteur exploitation sont de plusieurs types.

2.1 Démarche qualité ITIL[5]

En janvier 2005, compte tenu de la montée en puissance du système d'information du CHU, la Direction du Système d'Information s'est engagée dans une démarche qualité s'appuyant sur le référentiel des bonnes pratique ITIL. Ce référentiel organise la soutien et la fourniture de services informatiques autour de processus.

Après une première étape d'audit de maturité réalisé en 2005, il a été décidé de mettre en œuvre cinq processus ITIL :

Processus	Objectif des processus
La gestion des configurations	Fournir un modèle logique de l'infrastructure liée aux technologie de l'information
La gestion des incidents	Rétablir les niveaux de services convenus aussi vite que possible.
La gestion des problèmes	Minimiser les répercussions sur l'hôpital des incidents et problèmes informatiques
La gestion des changements	S'assurer que les méthodes et les procédures standardisées sont utilisées pour gérer efficacement et rapidement le changement.
La gestion des niveaux de services	Maintenir et améliorer la qualité de services informatiques.

2.2 Le centre de services

C'est le point de contact unique pour l'assistance à l'informatique et à la téléphonie. Il assure le support de niveau 1 pour l'ensemble des services applicatifs et le suivi de tous les incidents jusqu'à leur résolution. Il assure également le premier niveau de suivi de l'ensemble de la production et la surveillance des serveurs et processus.

Le centre de services traite en moyenne 1500 appels par mois pour notamment du dépannage et de l'assistance (à distance) pour les postes de travail informatiques, la téléphonie et les logiciels informatiques.

En dehors des heures d'ouverture du Centre de Services (jours ouvrables de 6h30 à 19h) l'astreinte informatique peut-être appelée. Elle a pour mission d'assurer un niveau de service minimum du système d'information permettant la continuité des applications critiques indispensables au bon fonctionnement du CHU. Le Personnel d'astreinte a pour mission de tout mettre en œuvre pour rétablir au plus tôt le bon fonctionnement du service dans le cadre des contrats de maintenance matériel et logiciel en cours.

[5] Information Technology Infrastructure Library

2.3 Intégration technique des applications

Elle consiste à intégrer dans l'environnement de production l'ensemble des logiciels applicatifs livrés et d'en assurer l'automatisation et le suivi de production.

Ce secteur gère plus de 80 services applicatifs, ce qui représente 8000 jobs automatisés répartis sur plus de 150 serveurs.

2.3.1 Automatisation et industrialisation de la production

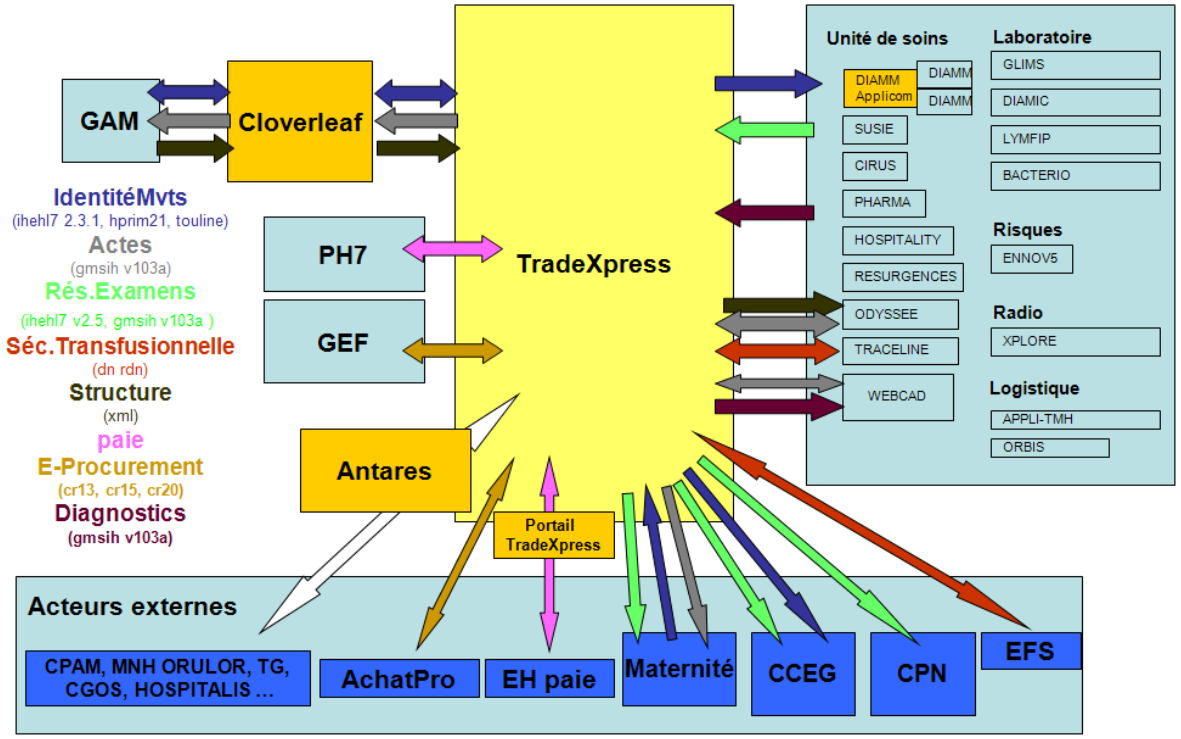
L'ordonnancement par un logiciel tel qu'OpconXps permet l'automatisation globale de la production. Ce ne sont ainsi pas moins de 2000 jobs qui s'exécutent chaque jour : backups de serveurs, interface entre base de données (échange de données entre deux bases),… Un tel produit a plusieurs avantages. Le plus évident est qu'il autorise une économie de personnel (personne n'est présent la nuit par exemple). De plus, il centralise les traitements ce qui permet un contrôle de l'ensemble de ces derniers depuis un seul point et une visualisation et une gestion des enchaînements et des événements. Depuis OpconXps, on accède donc aux jobs qui auraient pu tomber en erreur ou aux interdépendances entre plusieurs traitements (par exemple, un backup ne pourra s'exécuter que si l'arrêt de sa base s'est, au préalable, correctement déroulé). Enfin, il supporte tout type d'architecture (serveurs, PC) et tout système d'exploitation (Windows Server, Linux, Aix) en toute transparence ce qui apporte une grande flexibilité dans une structure d'une telle hétérogénéité.

2.3.2 Urbanisation

Un EAI[6] est un intergiciel qui automatise les échanges de données entre des applications hétérogènes d'un ou plusieurs systèmes d'information. Ce produit est indispensable dans la mesure où le CHU a à traiter une centaine de flux inter-applications. Prenons l'exemple de deux applications. D'une part, GAM (Gestion administrative du Malade), qui regroupe toutes les informations du malade : adresse, âge, numéro de sécurité sociale, l'attribution d'une chambre… D'autre part, l'application pharma, qui, elle, gère les prescriptions. Ainsi, elles vont pouvoir dialoguer entre elles grâce à l'EAI et il sera dès lors possible d'associer à un malade une prescription médicale.

Alors que l'EAI gère les flux internes, l'EDI[7] automatise les flux entre le CHU et les organismes externes. Par exemple, l'application Xplore gérant la dispensation des actes radiologiques, met à jour la GAM via l'EAI. Ces actes sont ensuite transmis à la CPAM (Caisse Primaire D'Assurance Maladie) via l'EDI afin d'être facturés.

[6] Enterprise Application Integration
[7] Échange de Données Informatisé



Synthèse des flux EAI et EDI

2.4 Télémaintenance

L'Exploitation met en place pour toute société extérieure ayant à intervenir pour des opérations de maintenance, d'assistance ou de dépannage sur des matériels ou des logiciels du CHU des accès par télémaintenance. L'accès se déroule en deux étapes :

- Création des identifiants personnels en lien avec l'équipe système pour chaque personne ayant à intervenir sur notre site après qu'elle ait signé un formulaire de confidentialité.
- Ouverture des accès aux différents serveurs en lien avec l'équipe réseau.

Ainsi, plus de 80 sociétés ont accès à la DSI par télémaintenance.

2.5 Salles informatiques

Le personnel de l'Exploitation a pour mission de gérer les salles informatiques. Cela entend la communication avec les services techniques (électricité, groupes froid, menuiserie,…) et l'installation de nouveaux matériels. Cette dernière fait l'objet d'une étude pour s'assurer que le nouvel équipement est compatible avec le plancher technique, les alimentations électriques, le système de climatisation et les connexions réseaux.

Les principaux matériels informatiques sont installés dans trois salles informatiques :

- Deux sur l'hôpital Marin qui hébergent le matériel du site de production
- Une à l'hôpital de Brabois adultes qui héberge le matériel du site de secours (quasiment saturée, la création d'une quatrième salle est en projet).

Les salles des hôpitaux de Brabois et Marin sont interconnectées par deux fibres optiques qui empruntent deux chemins différents afin d'éviter toute coupure, en cas de travaux sur la voierie par exemple.

L'accès aux salles de l'hôpital Marin est contrôlé par badges et est réservé au personnel de production. Tout intervenant extérieur désirant accéder aux salles doit être accompagné.

Ce contrôle des accès s'accompagne d'une protection électrique. L'ensemble de la distribution électrique du site de l'hôpital Marin est issu d'un seul poste de transformation sur le réseau EDF public commun aux usagers non prioritaires. Afin de sécuriser et d'assurer la

continuité de service des installations informatiques du CHU de Nancy, deux groupes électrogènes et quatre onduleurs apportent une sécurisation d'un niveau très élevé. De plus, chaque équipement sera double alimenté afin de s'affranchir d'une éventuelle panne d'alimentation.

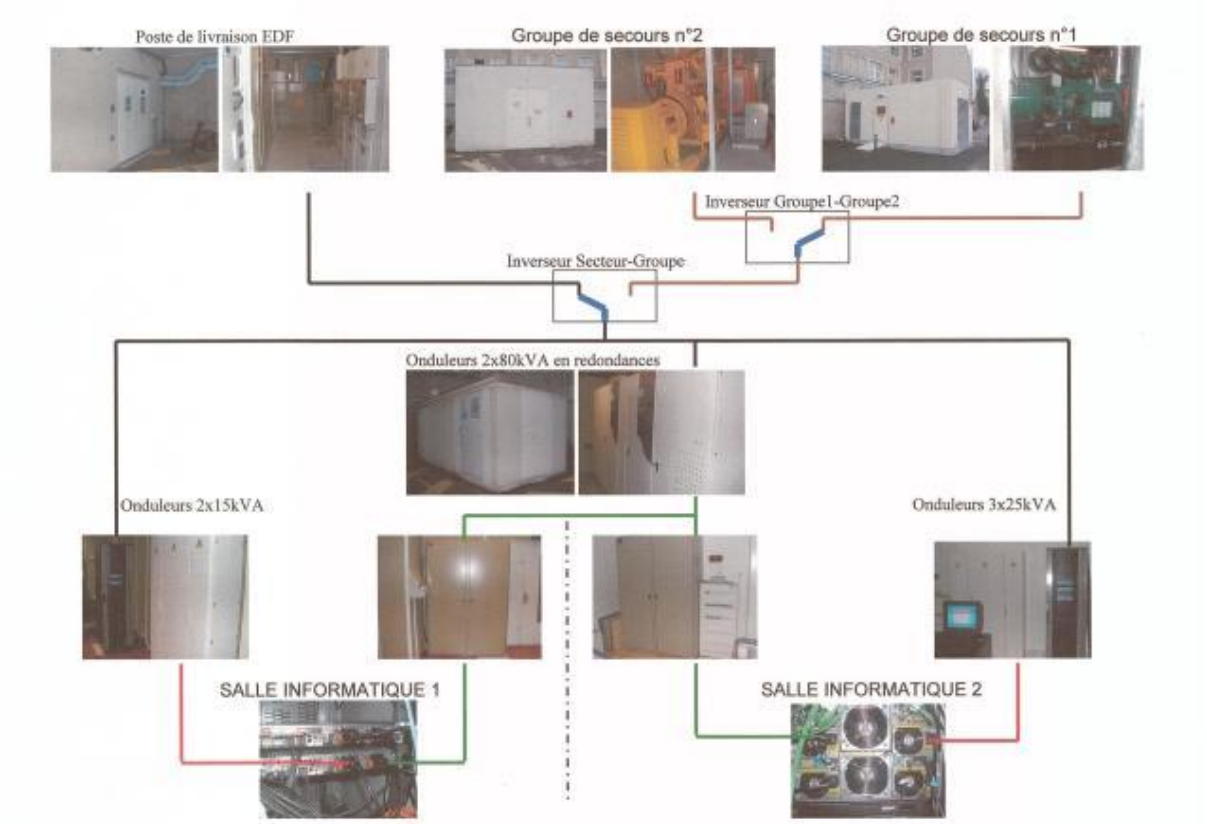


Schéma de l'installation électrique à Marin

2.6 Architecture de stockage

Le secteur Exploitation a également en charge l'architecture de stockage, point qui m'a tout particulièrement intéressé durant ce stage.

2.6.1 SAN[8]

Un réseau SAN est un réseau dédié à la mutualisation des ressources de stockage.

[8] Storage Area Network

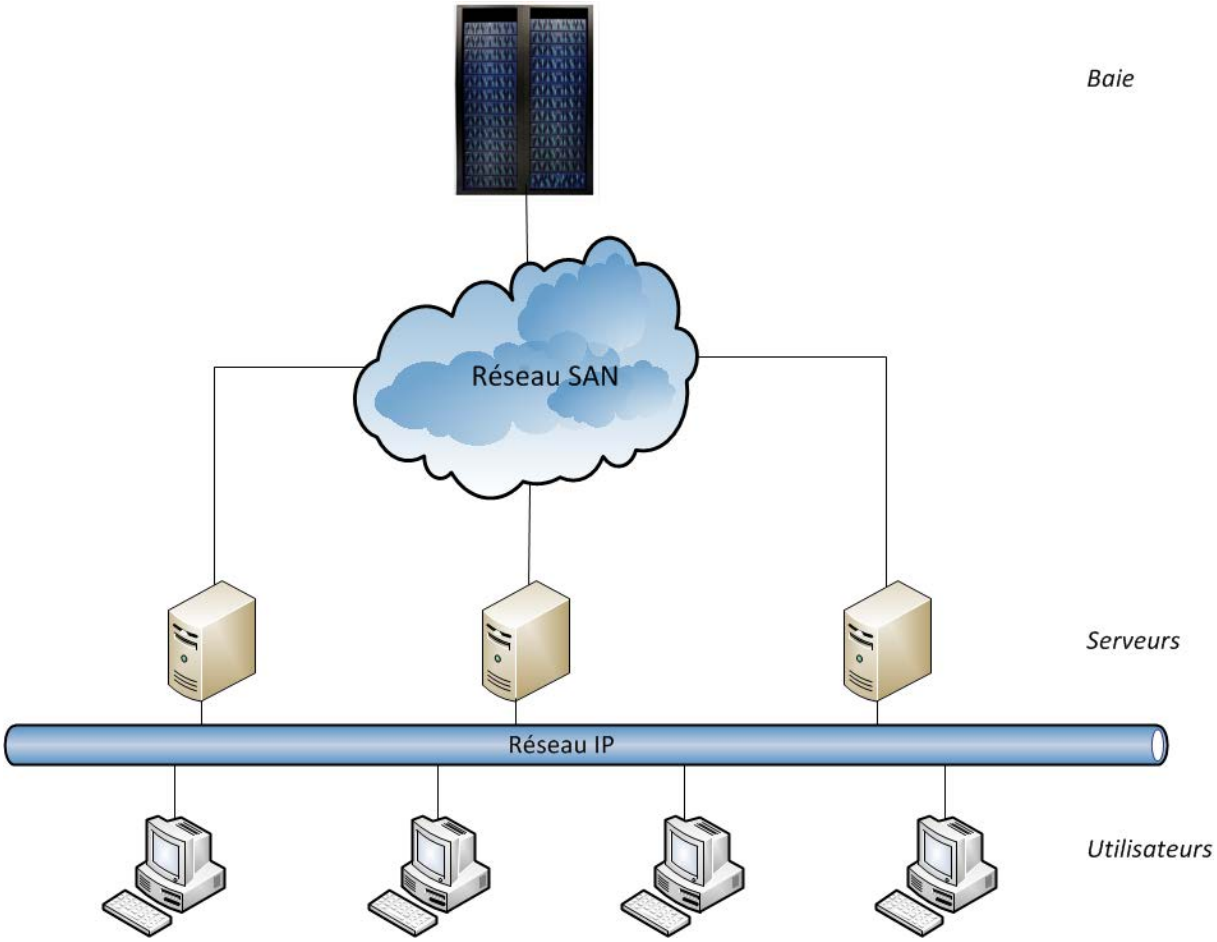


Schéma de principe d'une architecture SAN

Ce réseau s'appuie sur le protocole Fibre Channel qui permet d'interconnecter la baie de stockage aux serveurs à haut débit (8 Gb/s). Ces derniers vont donc avoir accès à cette volumétrie centralisée afin d'y stocker leurs applications et leurs bases de données. Les utilisateurs, quant à eux, accèdent aux serveurs et à leurs applications à travers le réseau IP.

Un réseau de ce type s'appuie sur une topologie Fabric. Ainsi, les nœuds réseau sont reliés entre eux via des switches. Elle se démarque de l'Ethernet, qui s'appuie sur le *broadcasting,* par le principe du *zoning.* Alors que le premier diffuse le message d'un émetteur vers l'ensemble des récepteurs, le second ne diffuse que vers un seul récepteur.

Dans le cas du CHU, il y a deux Fabrics s'étendant d'un site à l'autre. Chacun d'eux est constitué de deux switches optiques Brocade 4900 : un à Marin et un à Brabois.

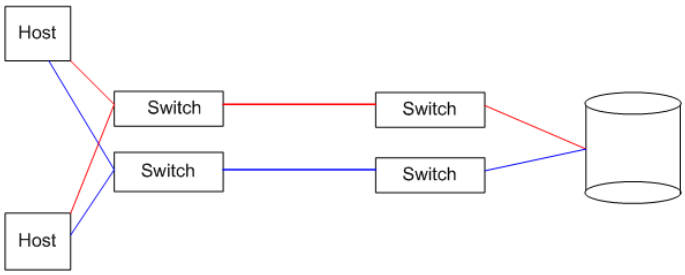


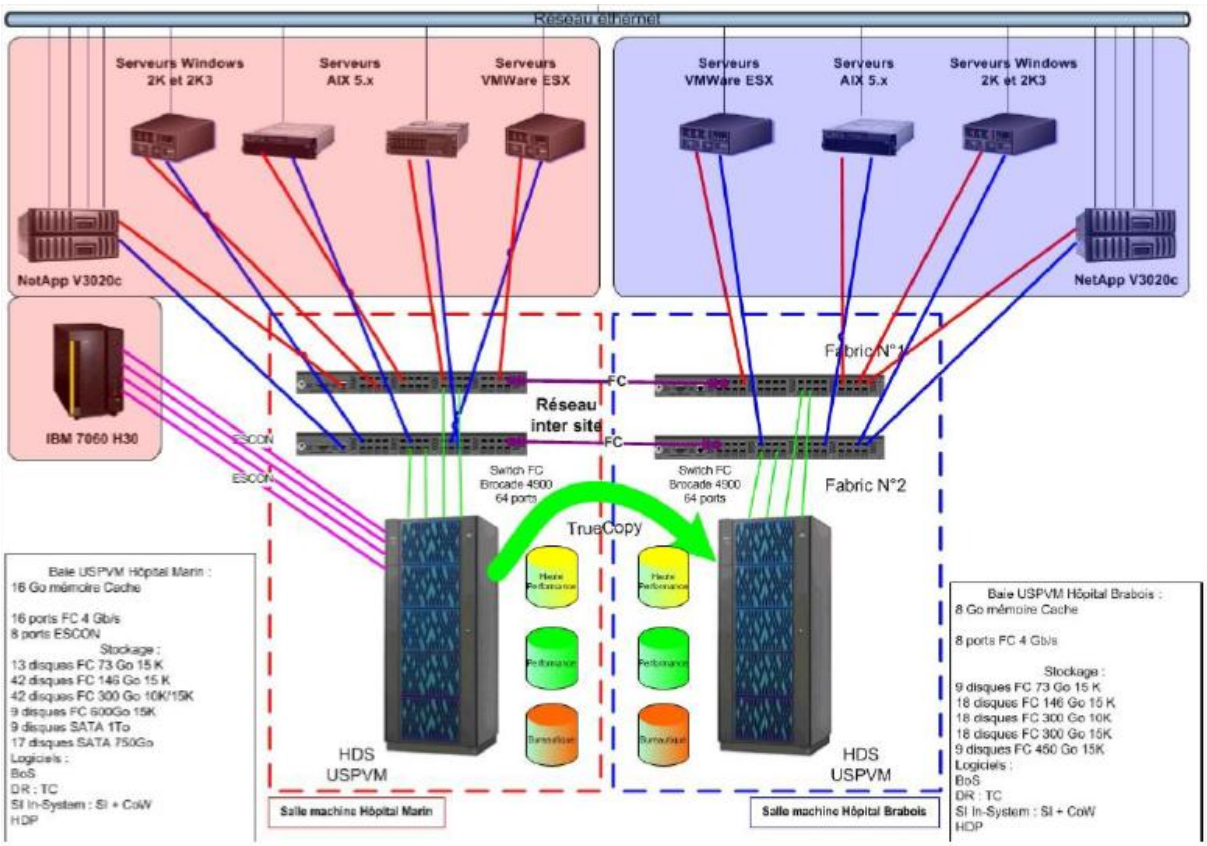
Schéma de principe de la topologie Fabric

Ces switches distants sont reliés entre eux grâce à un lien ISL[9] : une fibre monomode multiplexée à 2 Gb/s faisant transiter le protocole IP et Fibre Channel en parallèle. C'est grâce à

[9] Inter Site Link

elle que la réplication synchrone peut être mise en place. Ainsi, toutes les données sensibles stockées à Marin sont automatiquement transférées à Brabois. Cela permet en cas de sinistre sur le site principal de passer instantanément le site de secours comme site de production.

Plus précisément, le réseau SAN s'établit ainsi :



Réseau SAN du CHU (2010)

Les baies de stockage HDS[10] USPVM[11] disposent chacune de deux contrôleurs (8 ports chacun) totalement indépendants. Ces derniers permettent d'établir deux liens entre la baie et le serveur au travers des deux Fabrics. Cette redondance autorise donc la perte d'un contrôleur et/ou d'un switch tout en assurant la continuité de service.

Les baies disposent d'une volumétrie de 32 To[12] (27 à Marin et 5 à Brabois) répartie sur des disques de 73, 146, 300, 600, 750 et 1000 Go[13] à 15000 tours/min, le ratio prix/Go diminuant à mesure que la capacité du disque augmente. Cette différence de capacité varie en fonction des besoins en termes de performance. Par exemple, une application consommatrice d'I/O[14] ayant besoin d'une volumétrie de 146 Go aurait deux disques de 73 Go plutôt qu'un de 146 Go. Ainsi, pour ses lectures et écritures, elle disposerait de "2x15000 tours/min".

Une fois les disques physiques installés, il faut allouer de l'espace au serveur. Pour cela, on crée des LUNs[15] qui sont des partitions logiques des disques physiques. Ensuite, il convient de créer un Host Group par serveur. Il permet de rendre visible(s) uniquement la (les) LUN(s) qui lui est (sont) associée(s) au regard du serveur. Cette étape s'appelle le *mapping*.

[10] Hitachi Data Storage
[11] Universal Storage Platform Virtual Machine
[12] Téraoctets
[13] Gigaoctets
[14] Inputs/outpus : Entrées/Sorties
[15] Logical Unit Number

Dans une optique d'augmenter la sécurité et la performance, les disques sont organisés en RAID Groups de 8 disques selon la technique de RAID[16] 5. Cette dernière consiste à répartir des données sur 7 disques et à placer un calcul de parité sur un 8ème de manière circulaire. Ainsi, la parité (en rose ci-dessous), distribuée sur les 8 disques, permet en cas de perte d'un disque de reconstruire les données perdues. Cette sécurité accrue est couplée à un gain de performances. En effet, la LUN précédemment créée est attachée à un RAID Group et dispose donc de 8 axes de lecture/écriture.

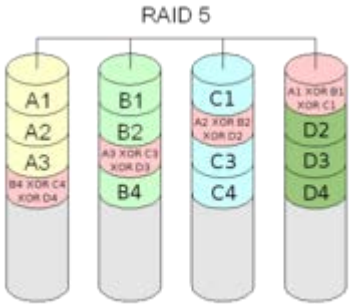


Schéma de principe du RAID 5

Une dernière étape est enfin nécessaire : le *zoning*. Il établit un chemin spécifique entre la baie et le serveur. Sans celui-ci, les deux équipements ne peuvent se "voir". On crée donc une zone à l'intérieur du switch sur chaque Fabric en mettant en lien le WWN[17] de la carte HBA[18] du serveur avec celui du port de la baie de stockage. Cette technique est également utilisée entre la baie de Marin et celle de Brabois pour permettre la réplication.

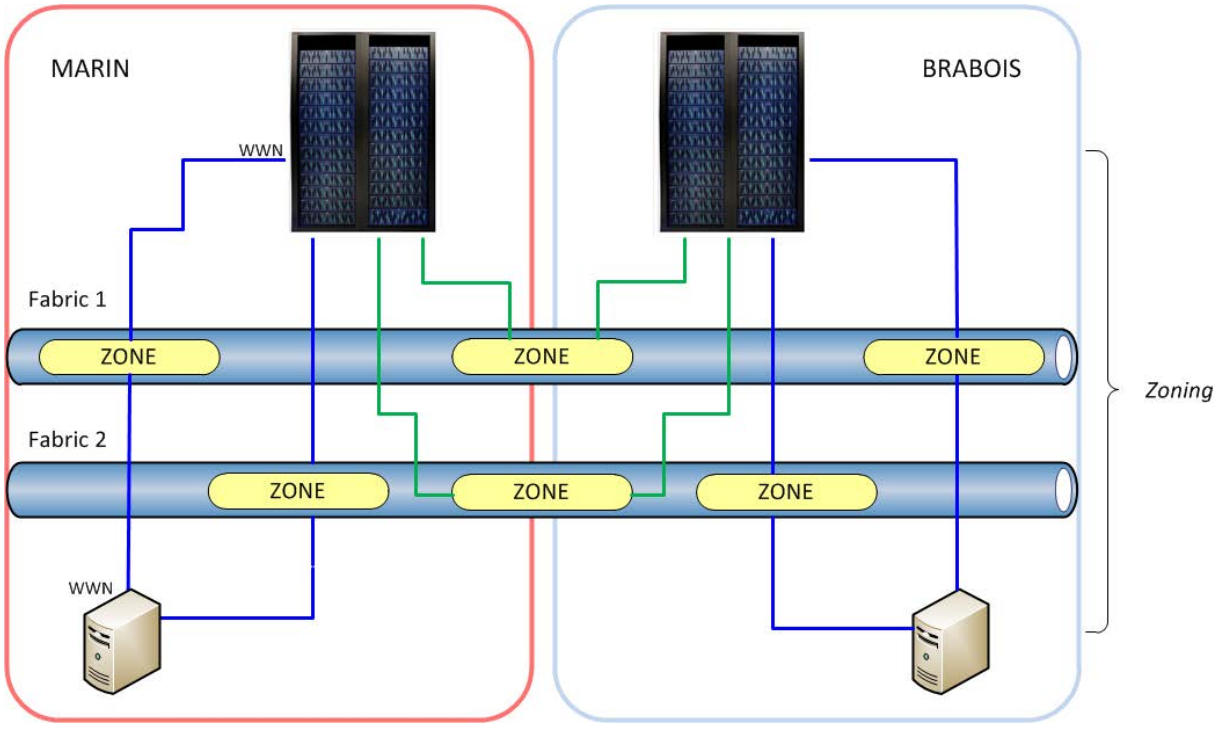


Schéma de principe du zoning

[16] Redundant Array of Independent Disks
[17] World Wide Name : Identifiant unique d'un équipement dans un réseau SAN (équivalent de l'adresse MAC)
[18] Host Bus Adapter : Carte d'extension permettant de connecter un serveur à un système de stockage

2.6.2 NAS[19]

Alors que le réseau SAN propose un accès bas niveau aux disques, le NAS met en avant des fichiers grâce à un protocole de partage de fichiers (CIFS[20] au CHU). Il est utilisé pour la bureautique, le partage de fichiers et permet de stocker de l'imagerie médicale.

Dans l'architecture en place, il y a convergence du SAN et du NAS dans la mesure où le NAS ne dispose pas de ses propres disques mais qu'il les partage avec ceux de la baie de stockage SAN. Il n'a donc ici qu'une mission de contrôleur.

2.7 Architecture de Sauvegarde

Par sauvegarde on entend le fait de « répliquer » tout ou partie des éléments du système d'information sur des supports externes dans l'optique de les reconstituer en cas de panne ou de dysfonctionnement (attaque virale, erreur de manipulation, sinistre). Ces sauvegardes multiplateformes sont gérées par le logiciel Netbackup en collaboration avec l'ordonnanceur Opcon qui, lui, va gérer leur fréquence.

Les données sont tout d'abord enregistrées sur un système de sauvegarde sur disque, des DXi Quantum disposant de 40 To de disques autorisant jusqu'à 20 sauvegardes simultanées. Elles subissent ensuite une déduplication, ce qui implique un découpage en blocs. Son objectif est de ne stocker qu'une seule fois un même tronçon. Aussi, une nouvelle occurrence d'un tronçon déjà présent n'est pas à nouveau sauvegardée, mais remplacée par un pointeur vers celui-ci. Actuellement, l'espace utilisé est divisé par 12, ce facteur augmentant à mesure que le nombre de blocs référence augmente.

Une fois les données factorisées, elles sont envoyées vers le DXi de Brabois afin de toujours disposer d'une deuxième copie sur un site distant, ceci afin d'échapper à un sinistre de site.

La sauvegarde se trouvant désormais sur chaque site, elle est dupliquée sur une librairie SCALAR. Cette dernière, contrairement au DXi, est constituée de bandes magnétiques de type LTO[21]5 et permet d'y stocker les backups critiques et/ou à longue rétention (jusqu'à 5 ans). Les bandes ne supportant pas la déduplication, les données sont réhydratées. On s'assure par la même occasion que l'on est bien capable de les reconstituer.

Autrefois, les données étaient directement sauvegardées sur la librairie mais cette technique présentait plusieurs inconvénients. Tout d'abord, le nombre de sauvegardes simultanées est limité par rapport à une sauvegarde sur disques. De plus, le temps d'accès à un backup est bien plus lent car il faut le temps que la bande soit montée dans le lecteur et que ce dernier la déroule jusqu'au bon endroit. Enfin, les interventions manuelles étaient nombreuses afin de sortir les bandes pleines du robot et de les changer par des vides.

[19] Network Attached Storage
[20] Common Internet File System
[21] Linear Tape-Open

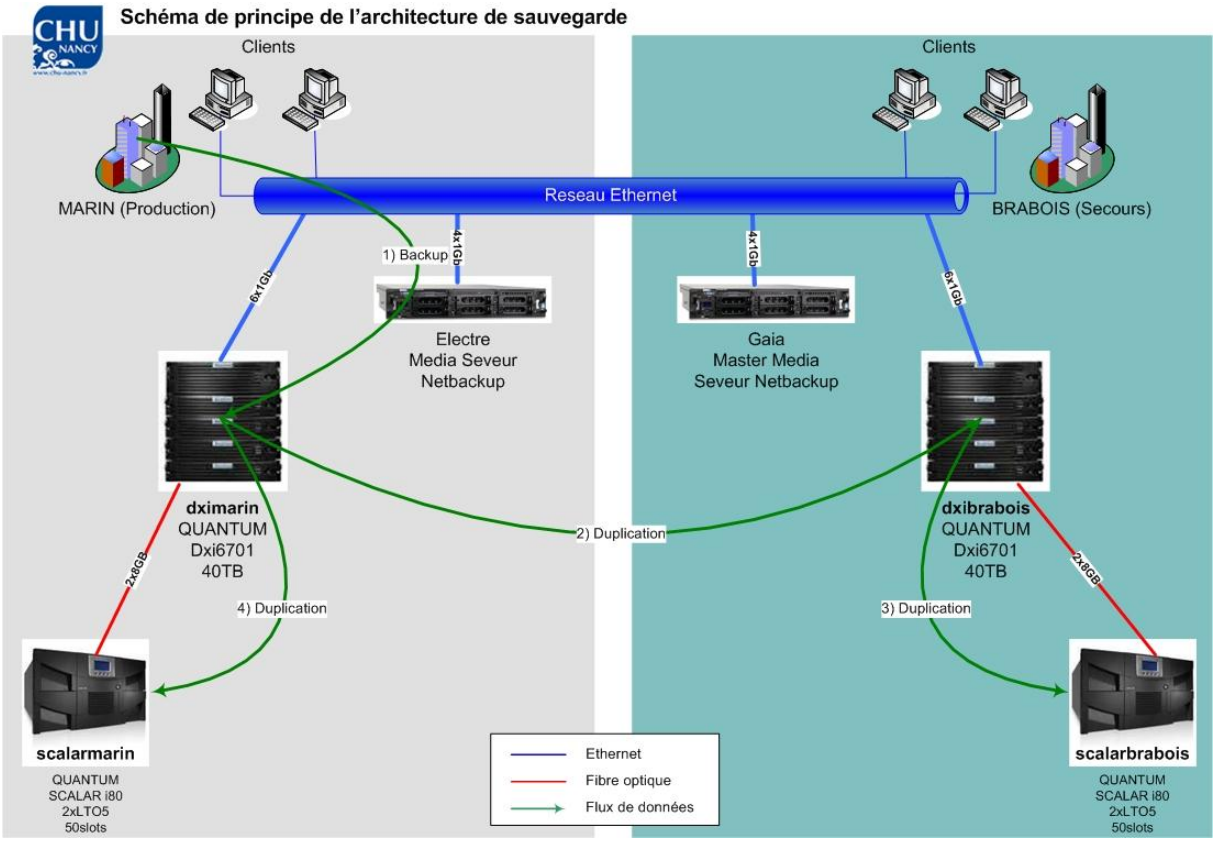


Schéma de principe de l'architecture de sauvegarde

3. Le projet

3.1 Présentation

3.1.1 Sujet de stage

Le stage consiste à la mise en œuvre du nouveau système de stockage SAN du CHU de NANCY.

Il s'agit de mettre en place de nouvelles baies sur les sites de production et de secours, puis de migrer les données des anciennes vers les nouvelles baies tout en maintenant en condition opérationnelle les systèmes informatiques.

3.1.2 Contexte

Ce projet de migration du système de stockage s'inscrit dans un contexte d'évolution perpétuelle de la volumétrie, de l'ordre de 15% par an. Mi-2011, alors que les besoins en termes d'espace se faisaient de plus en plus ressentir, la question sur l'évolution du SAN s'est donc tout naturellement posée.

Un simple ajout de disque dans la baie USPVM aurait pu être envisagé si le contrat de maintenance n'arrivait pas à échéance en Juin 2012. En effet, la reconduction de celui-ci aurait été très onéreuse, plus encore que l'achat d'une nouvelle baie… Il a donc été décidé d'installer une nouvelle architecture de stockage.

L'USPVM est une technologie haut de gamme qui s'imposait par le Mainframe. En effet, cet ordinateur central d'une grande puissance dispose de ports ESCON, seulement disponibles sur l'USP. Or, il a été progressivement remplacé par des serveurs, à tel point qu'en Juin 2010, il a été totalement arrêté. Une telle technologie de stockage, jusque-là indispensable, était donc devenue superflue.

3.1.3 Nouvelle configuration matérielle

Le choix s'est finalement porté sur une baie HDS AMS[22] 2500. C'est une technologie moyenne gamme - qui entraîne donc une baisse des coûts - et qui reste du constructeur HDS. Ceci est très important car le changement de constructeur engage une rupture technologique. Le personnel reste donc opérationnel sans engendrer de coûts supplémentaires pour le reformer. En outre, elle propose un équilibrage dynamique de charge entre ses deux contrôleurs. Ainsi, la baie va répartir la charge automatiquement si un contrôleur devient trop sollicité (+ de 70% d'utilisation du processeur). Les temps de réponse globale s'en voient dès lors améliorés en évitant tout goulot d'étranglement. Enfin, l'AMS ajoute un niveau d'agrégation supplémentaire. En effet, les RAID Groups sont désormais regroupés au sein d'un pool. Ainsi, les LUNs ne sont plus réparties sur un RAID Group (8 disques) mais sur, potentiellement, l'ensemble des disques de la baie (jusqu'à 195) - ce qui en fait autant d'axes de lecture/écriture -, d'où une augmentation des performances.

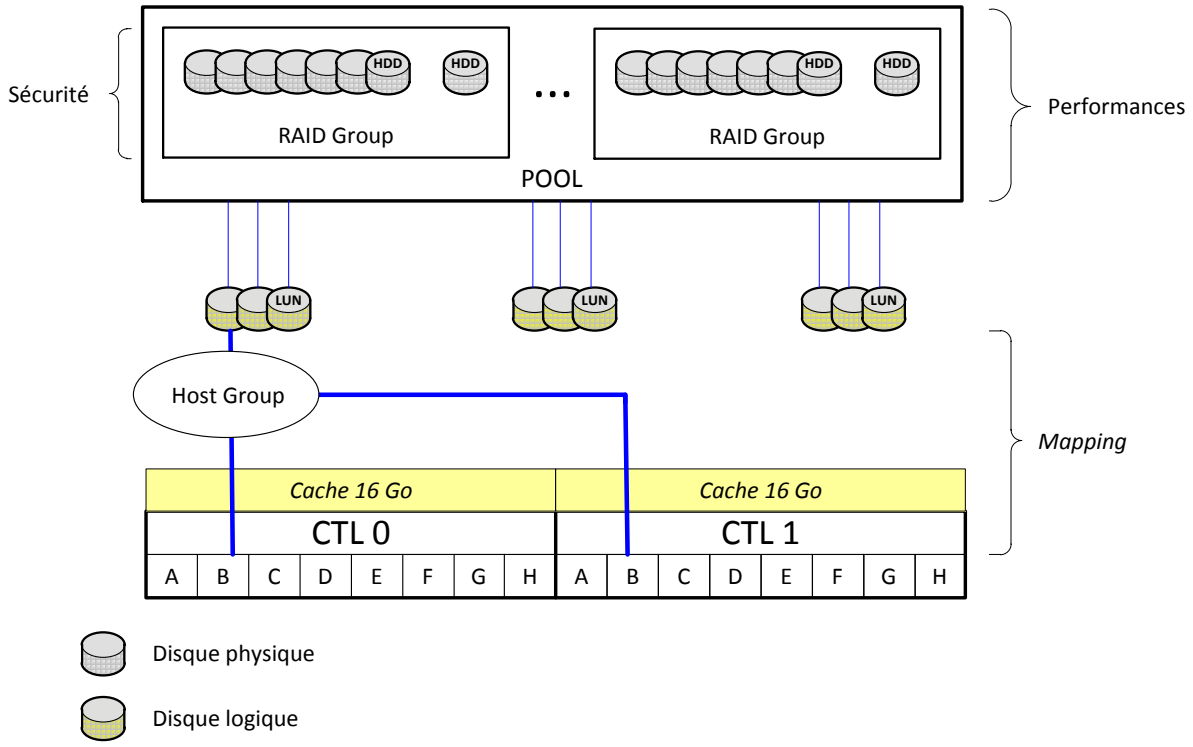


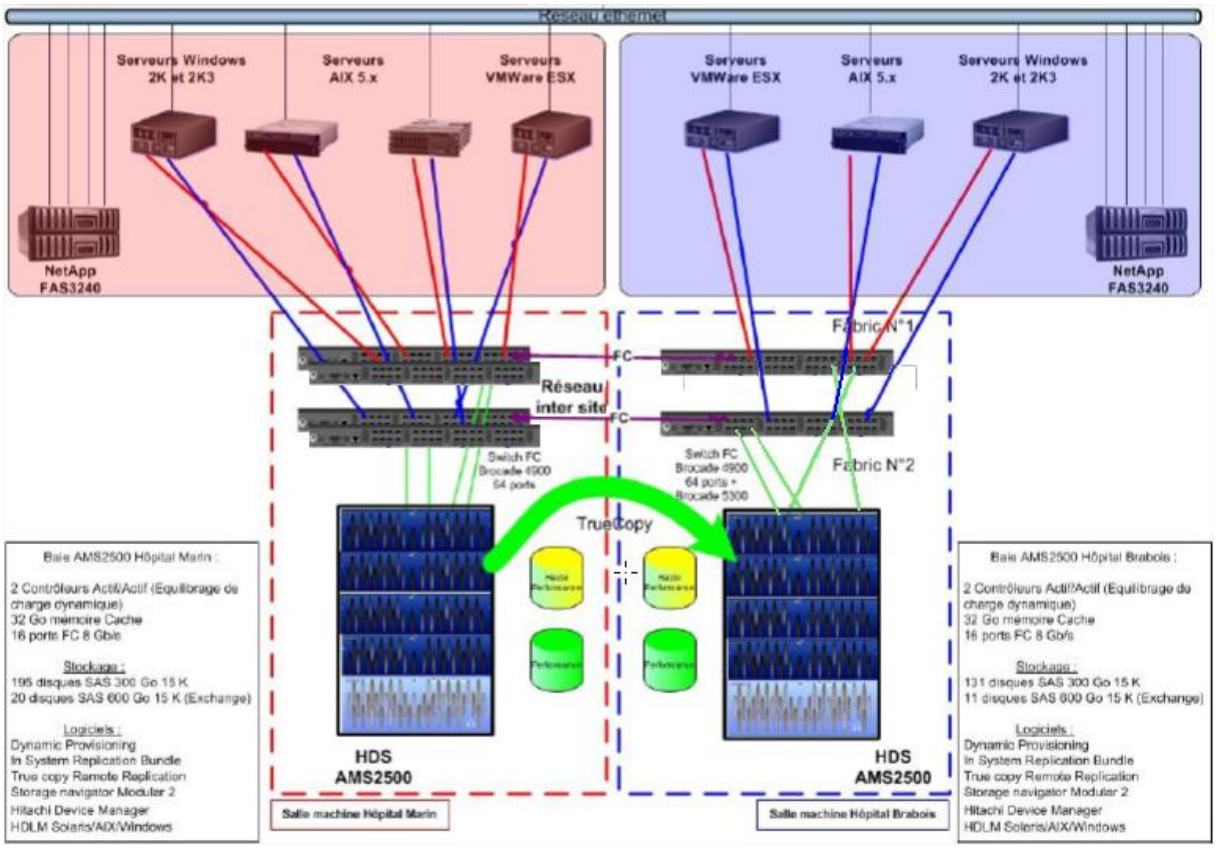
Schéma de principe de la baie

La volumétrie a été pensée pour quatre ans en partant d'une estimation d'augmentation annuelle de 15%. On disposait de 32 To sur l'ancienne baie, ce qui nous amène à 56 To au bout de quatre ans. A cela, il faut rajouter 18 To pour la messagerie et une vingtaine de To pour la marge de manœuvre. On dispose donc d'une centaine de To, environ 60 à Marin et 40 à Brabois. La messagerie, peu consommatrice d'I/O sera placée sur des disques de 600 Go (Pool 0), le reste étant disposé sur des disques de 300 Go (Pool 1).

De plus, afin de contenir l'évolution du nombre de serveurs, un nouveau switch Brocade 5300 a été ajouté sur chacun des fabrics, offrant 64 ports supplémentaires.

Enfin, le NAS, qui convergeait jusque-là avec le SAN, dispose désormais de ses propres disques (40 To). Ce changement s'explique par le fait que la convergence n'est plus aussi intéressante financièrement qu'elle pouvait l'être à l'époque de l'USPVM. De plus, cela peut entraîner une panne en série car si le SAN tombe, le NAS tombe aussi.

[22] Adaptable Modular Storage



Nouvelle architecture de stockage

3.1.4 Etapes de la migration

En amont de la migration, plusieurs étapes sont nécessaires. Tout d'abord, un inventaire complet des serveurs connectés au SAN doit être fait. Ensuite, on contrôle la compatibilité des OS[23], firmwares et cartes HBA avec la nouvelle baie. Ceci fait, il faut définir les méthodes de migration, celles-ci étant propres à l'OS du serveur. De plus, elles doivent privilégier une continuité de service, point très important dans un CHU. En effet, une migration à chaud est totalement transparente pour l'utilisateur alors qu'une à froid entraîne un arrêt total du serveur de production.

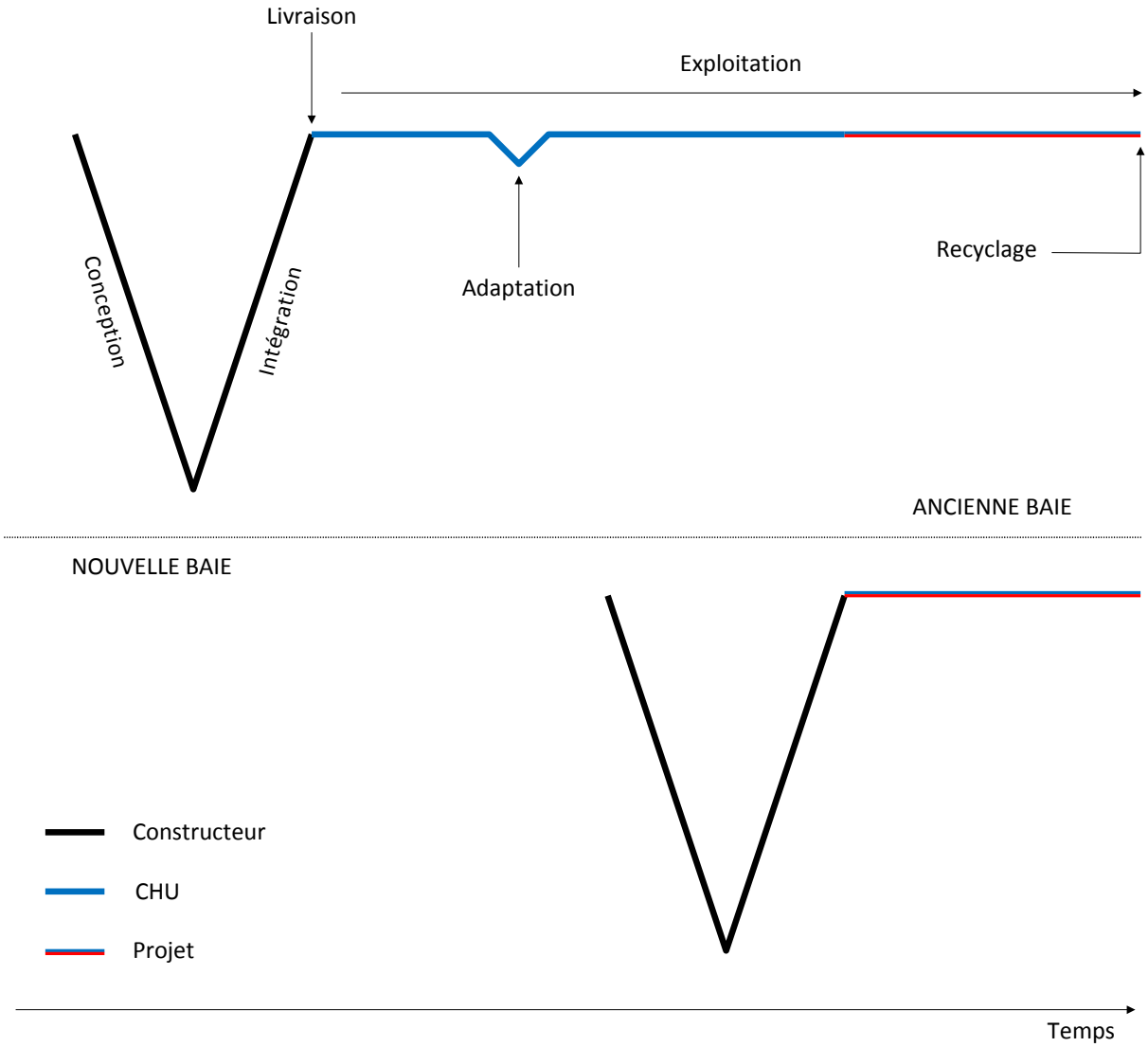
La nouvelle architecture de stockage peut dès lors être installée afin de tester les méthodes de migration sur des environnements dits de pré-production. Ces derniers sont des copies conformes des machines de production et permettent, ici, de simuler une migration. En d'autres occasions, ils servent également à tester si une mise à jour d'application se passe sans accroc. Je n'ai assisté à ces étapes de préparation qu'en tant que spectateur.

En revanche, j'ai activement participé au cœur du projet : la migration. Car une fois les méthodes validées, il convient de les mettre en pratique. Pour cela, une fois la volumétrie définie, il faut créer la LUN sur la nouvelle baie. Il faut ensuite créer le Host Group, mapper ladite LUN et créer le zoning.

A ce stade, le serveur dispose d'un espace de stockage auquel lui seul peut accéder. Par la suite, si l'application est sensible, la réplication inter-sites doit être mise en place. Pour ce faire, on crée sur la baie de Brabois une LUN strictement identique que l'on appaire à celle de Marin : toute donnée inscrite à Marin est immédiatement inscrite à Brabois. Dès lors, on peut migrer les données de l'ancienne vers la nouvelle baie à Marin, celles-ci seront automatiquement retranscrites à Brabois de manière synchrone.

Enfin, dès que la migration est validée, on déconnecte le serveur de l'ancienne baie.

[23] Operating System



Cycle de vie des baies de stockage

3.2 Gestion de projet

3.2.1 Les besoins client

Afin de mener à bien ce projet, mon tuteur a exprimé plusieurs besoins. Tout d'abord, je devais mettre au point un outil d'aide à la création de LUNs que j'utiliserai au quotidien lors de la migration et qui servira durant toute l'exploitation de la baie après mon départ. Dans un premier temps, il fallait qu'il permette d'industrialiser le processus de création des LUNs, c'est-à-dire l'automatiser en réduisant au maximum voire en éliminant totalement toute intervention manuelle directement sur la baie de stockage. Deuxièmement, il se devait d'assurer une traçabilité de l'ensemble des actions que nous avons pu mener. Enfin, grâce à lui, nous devions pouvoir suivre l'évolution de la volumétrie allouée.

En parallèle, il fallait s'assurer que les performances des équipements étaient conformes aux données constructeur. Cette métrologie devait se faire au moyen d'un logiciel libre et l'ensemble du personnel du secteur Exploitation devait pouvoir y accéder par un point unique.

Un serveur lui serait dédié tant que son OS soit libre, compatible avec l'environnement de travail et qu'il permette la création de scripts Shell.

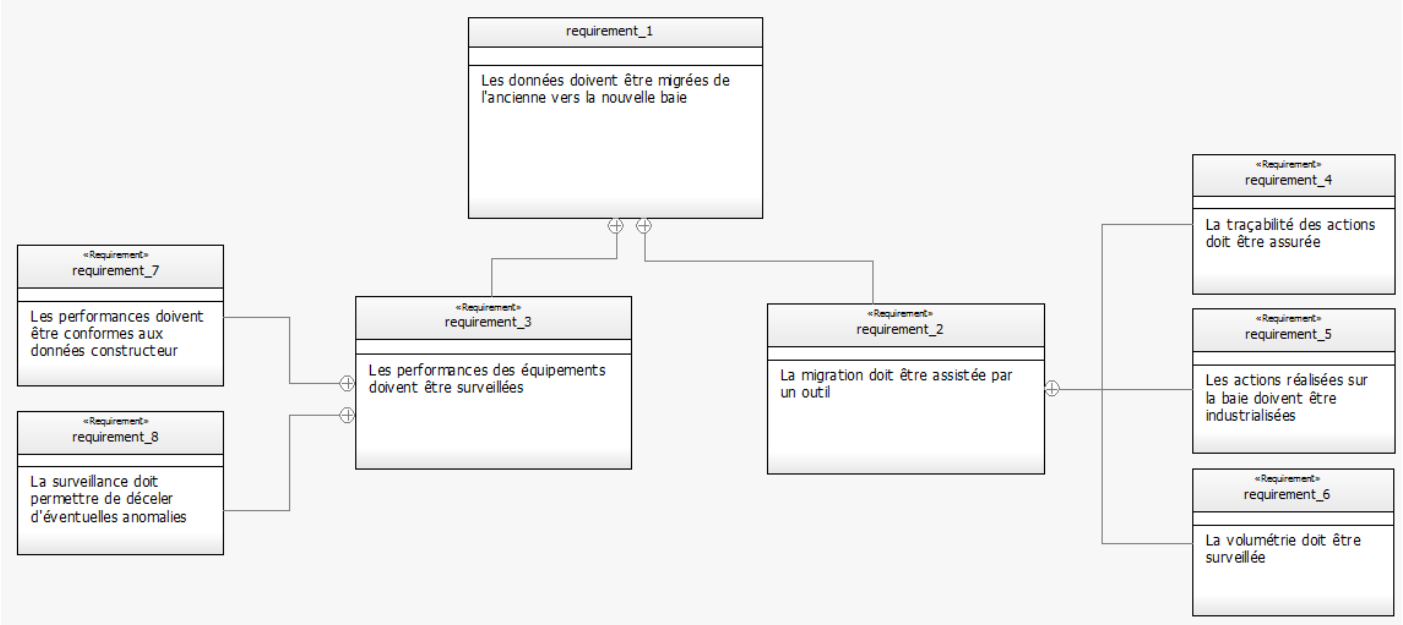
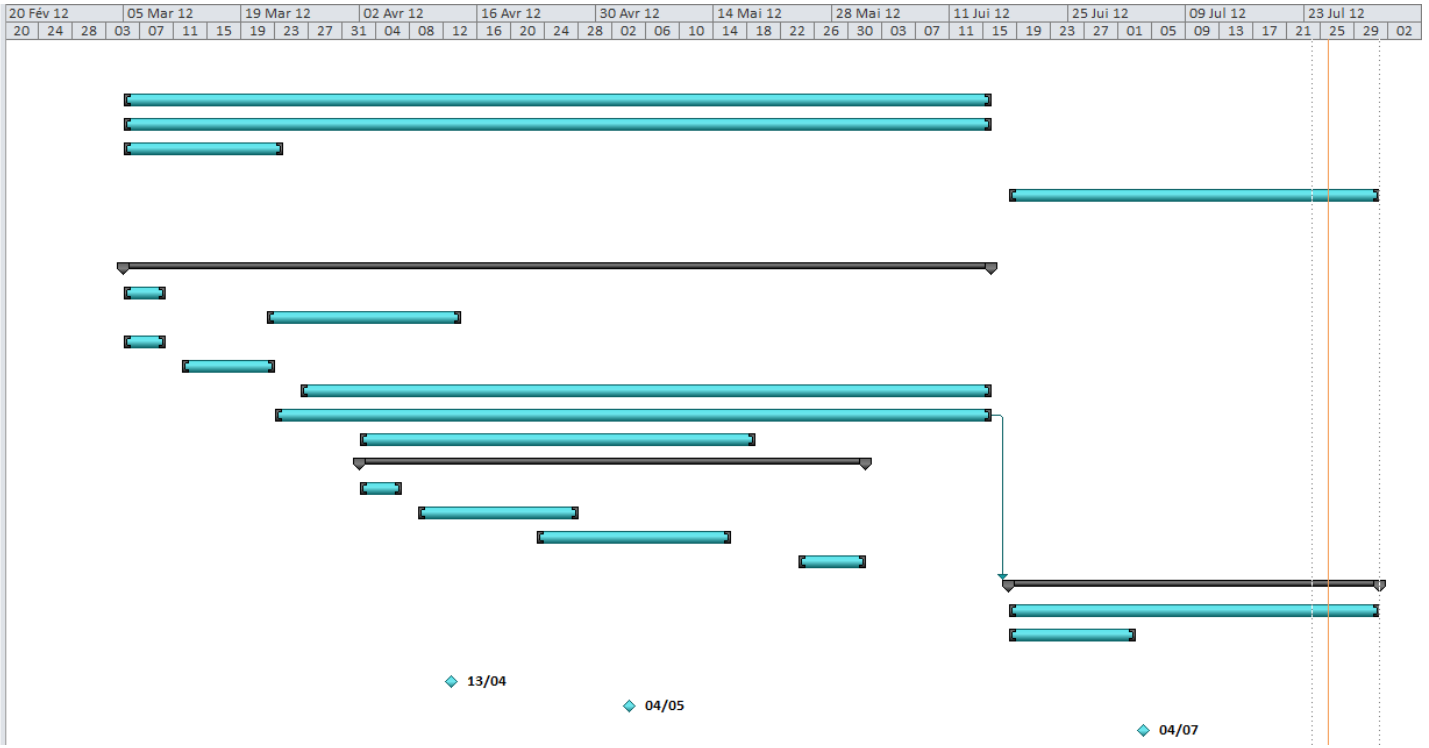


Diagramme d'exigences

D'un point de vue temporel, deux contraintes m'ont été fixées : l'outil d'aide à la création de LUNs devait être exploitable le 23 Mars, pour le début de la migration. Deuxièmement, la métrologie devait être opérationnelle au 15 Juin car la migration d'une grosse application était prévue ce jour-ci. Nous devions donc pouvoir surveiller si la baie tenait la charge supplémentaire qu'on lui imposait. Il est à noter enfin que la migration devait être terminée au 15 Juin, date à laquelle la maintenance de l'ancienne baie USPVM se terminait.

Ces contraintes temporelles ont toutes été respectées. En effet, l'outil était prêt le 21 Mars mais j'ai continué à l'optimiser jusqu'au 02 Juillet. J'ai complété ce dernier par une méthodologie de son utilisation. La métrologie, quant à elle, s'est achevée le 31 Mai et la migration le 15 Juin.

Nom de la tâche	Durée	Début	Fin
Planning prévisionnel			
Migration	75 jours	Lun 05/03/12	Ven 15/06/12
Métrologie	75 jours	Lun 05/03/12	Ven 15/06/12
Mise au point de l'outil de création de LUNs	15 jours	Lun 05/03/12	Ven 23/03/12
Production	32 jours	Lun 18/06/12	Mar 31/07/12
Planning réel			
Migration	**75 jours**	**Lun 05/03/12**	**Sam 16/06/12**
Définition des méthodes de migration	5 jours	Lun 05/03/12	Ven 09/03/12
Test des méthodes de migration	17 jours	Jeu 22/03/12	Ven 13/04/12
Définition des besoins	5 jours	Lun 05/03/12	Ven 09/03/12
Mise au point de l'outil de création de LUNs	9 jours	Lun 12/03/12	Jeu 22/03/12
Optimisation de l'outil de création de LUNs	60 jours	Lun 26/03/12	Ven 15/06/12
Migration serveurs Aix	61 jours	Ven 23/03/12	Ven 15/06/12
Migration serveurs Windows	35 jours	Lun 02/04/12	Ven 18/05/12
Métrologie	**44 jours**	**Lun 02/04/12**	**Jeu 31/05/12**
Définition des besoins	5 jours	Lun 02/04/12	Ven 06/04/12
Installation et configuration	15 jours	Lun 09/04/12	Ven 27/04/12
Création scripts	17 jours	Lun 23/04/12	Mar 15/05/12
Mise à jour Cacti	6 jours	Jeu 24/05/12	Jeu 31/05/12
Production	**32 jours**	**Lun 18/06/12**	**Mar 31/07/12**
Production	32 jours	Lun 18/06/12	Mar 31/07/12
Optimisation de l'outil de création de LUNs	11 jours	Lun 18/06/12	Lun 02/07/12
Réunion migration n°1	0 jour	Ven 13/04/12	Ven 13/04/12
Réunion migration n°2	0 jour	Ven 04/05/12	Ven 04/05/12
Visite tuteur universitaire	0 jour	Mer 04/07/12	Mer 04/07/12



Planning réel/Planning prévisionnel

3.2.2 Les solutions choisies

Afin de répondre à ces exigences, j'ai dû faire des choix techniques. Tout d'abord, concernant l'outil de création de LUNs, la traçabilité et le suivi de la volumétrie m'ont très vite conduit à un tableur. En effet, il dispose de tous les outils nécessaires pour tracer dynamiquement un graphique et son affichage sous forme de tableau est parfait pour archiver nos actions. J'avais donc le choix entre Calc d'Openoffice et Excel de Microsoft. Il m'a semblé pertinent de choisir le second, car bien que payant, il est installé en natif sur l'ensemble des postes du CHU.

A l'heure de suivre les performances des équipements du SAN, ma réflexion s'est portée sur Cacti et Centreon. En effet, tous deux sont des logiciels libres de mesure de performances réseau et serveur. J'ai finalement opté pour Cacti qui jouit d'une grande communauté. Les mises à jour sont donc fréquentes et toutes mes questions peuvent trouver une réponse sur les forums.

Cacti n'est pas limitant en termes d'OS et peut s'installer aussi bien sur du Windows que du Linux. Cependant, il était exigé qu'il soit libre, ce qui élimine d'office Windows (non-libre et payant) et RedHat (libre mais payant). De plus, il doit être compatible avec l'ordonnanceur (Opcon) et le logiciel de sauvegarde (Netbackup) : Fedora ne fait donc pas l'affaire. J'ai donc choisi Debian qui répond à tous ces critères.

Afin de pouvoir accéder à ces outils par un point unique, il fallait mettre en place un système de gestion de contenu qui permet, entre autre, de concevoir des portails Web. Il se devait d'être libre et facile d'accès, n'ayant aucune connaissance dans ce domaine. Joomla ! a donc été adopté.

Enfin, cet ensemble se devait de tourner sur un serveur dédié. Parmi les serveurs inactifs, le HP Proliant DL140 m'a semblé être le choix le plus judicieux. Ce serveur moyenne gamme propose en effet des ressources supérieures à ce que peut exiger Cacti et lui offre donc des perspectives d'évolution. Sa capacité de 80 Go est en outre tout à fait suffisante quant aux besoins du logiciel de mesures.

3.3 Travaux réalisés

3.3.1 Outil d'aide à la création de LUNs

Cet outil permet tout d'abord de garder une trace de l'ensemble des LUNs que nous avons à créer. Pour cela, dès que le secteur Systèmes nous envoie une demande de changement (selon la procédure ITIL), il convient de remplir le tableau notamment avec la taille de la partition, son pool, le hostgroup qui y aura accès et par quel port et enfin les informations nécessaires à la réplication.

POO	LUN	Capacity MB	Nombre de blocks	Port 1	Port 2	Hosts	HLUN	hdisk	VG	Point de montage ou seveurs ESX	Réplication dev_group	Réplication dev_name	Acc. Wide Str. M...	Pool 0 (MB)	Pool 0 réel (MB)	Modifié par	Date
000	0123	40 527	82 999 296	0 C	1 C	Mintaka	0000	hdisk23	rootvg				oui	40 527	49 152	JE. SIMONIN	6/4/12
000	0124	70 000	143 360 000	0 C	1 C	Mintaka	0001	hdisk24	vgh00	/soft /export /archIGPRD /redologs /expl /archBMACPRD /archANAP	mintaka	P012480066	oui	70 000	73 728	JE. SIMONIN	6/4/12
000	0125	150 000	307 200 000	0 C	1 C	Mintaka	0002	hdisk25	vgh01	/igrec /igfor /igprd	igestor	P012580067	oui	150 000	172 032	JE. SIMONIN	6/4/12
000	0126	7 000	14 336 000	0 C	1 C	Mintaka	0003	hdisk26	vgh03	/divat /archdivat /archbackup	divat	P012680068	oui	7 000	24 576	JE. SIMONIN	6/4/12
000	0127	6 000	12 288 000	0 C	1 C	Mintaka	0004	hdisk27	vgh04	/vaxtst /vaxprd	infovax	P012780069	oui	6 000	24 576	JE. SIMONIN	6/4/12

Traçabilité de la création de LUNs

Une fois ces données saisies, on a donc une génération automatique des lignes de commandes nécessaires à la création, au mapping et à la réplication de LUNs.

Création LUN	Mapping LUN Port 1	Mapping LUN Port 2	horcm.10				
Echo y | auluadd -unit M_AMS -lu 123 -dppoolno 0 -size 40527m -widestriping enable -fullcapacity	Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 0 -lu 12	Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 0 -lu 12					
Echo y | auluadd -unit M_AMS -lu 124 -dppoolno 0 -size 70000m -widestriping enable -fullcapacity	Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 1 -lu 12	Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 1 -lu 12	mintaka	P0124S0066	CL1-C	0	0124
Echo y | auluadd -unit M_AMS -lu 125 -dppoolno 0 -size 150000m -widestriping enable -fullcapacity	Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 2 -lu 12	Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 2 -lu 12	igestor	P0125S0067	CL1-C	0	0125
Echo y | auluadd -unit M_AMS -lu 126 -dppoolno 0 -size 7000m -widestriping enable -fullcapacity e	Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 3 -lu 12	Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 3 -lu 12	divat	P0126S0068	CL1-C	0	0126
Echo y | auluadd -unit M_AMS -lu 127 -dppoolno 0 -size 6000m -widestriping enable -fullcapacity e	Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 4 -lu 12	Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 4 -lu 12	infovax	P0127S0069	CL1-C	0	0127

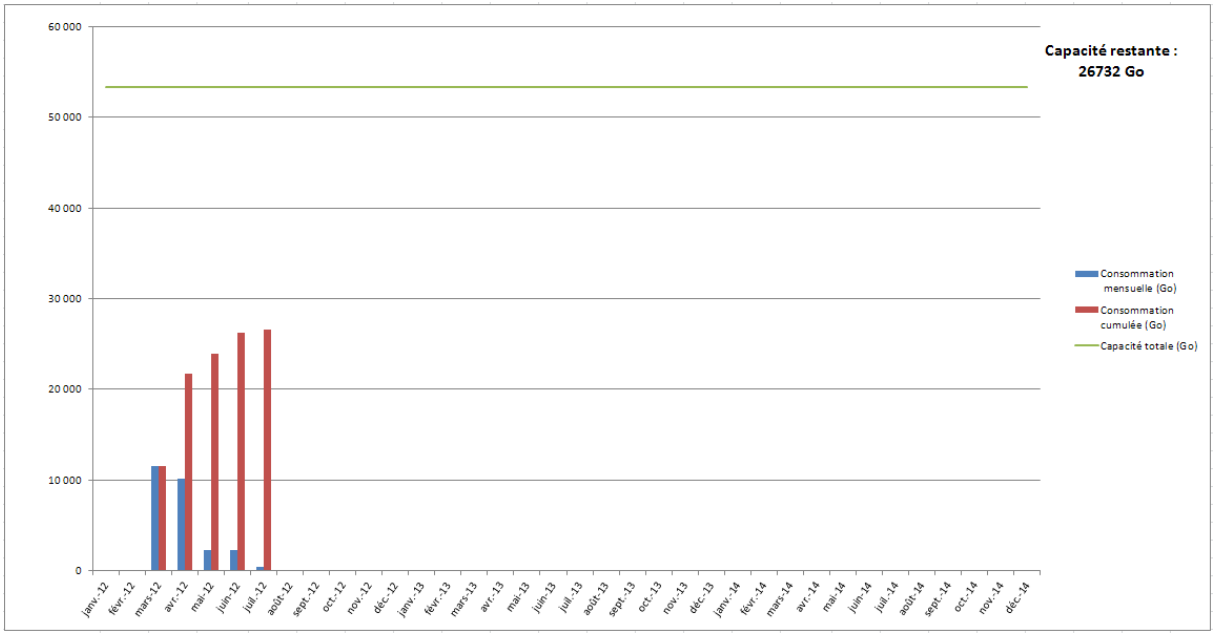
Lignes de commandes de création de LUNs

Il ne reste plus qu'à copier ces commandes et les envoyer sur la baie directement. Nous profitons donc d'une étape indispensable qu'est la documentation des interventions effectuées sur la baie pour automatiser le processus. Ainsi, nous avons contourné l'interface graphique qui ne permet aucune industrialisation et qui peut être source d'erreurs.

```
SNM2>Echo y | auluadd -unit M_AMS -lu 123 -dppoolno 0 -size 40527m -widestriping enable -fullcapacity enable
Are you sure you want to set the logical unit? (y/n [n]): The logical unit has been set successfully.

SNM2>Echo y | auluadd -unit M_AMS -lu 124 -dppoolno 0 -size 70000m -widestriping enable -fullcapacity enable
Are you sure you want to set the logical unit? (y/n [n]): The logical unit has been set successfully.

SNM2>Echo y | auluadd -unit M_AMS -lu 125 -dppoolno 0 -size 150000m -widestriping enable -fullcapacity enable
Are you sure you want to set the logical unit? (y/n [n]): The logical unit has been set successfully.

SNM2>Echo y | auluadd -unit M_AMS -lu 126 -dppoolno 0 -size 7000m -widestriping enable -fullcapacity enable
Are you sure you want to set the logical unit? (y/n [n]): The logical unit has been set successfully.

SNM2>Echo y | auluadd -unit M_AMS -lu 127 -dppoolno 0 -size 6000m -widestriping enable -fullcapacity enable
Are you sure you want to set the logical unit? (y/n [n]): The logical unit has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 0 -lu 123
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 1 -lu 124
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 2 -lu 125
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 3 -lu 126
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 4 -lu 127
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 1 -lu 124
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 2 -lu 125
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 3 -lu 126
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 0 C -gname Mintaka -hlu 4 -lu 127
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 0 -lu 123
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 1 -lu 124
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 2 -lu 125
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 3 -lu 126
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit M_AMS -add 1 C -gname Mintaka -hlu 4 -lu 127
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.
```

Insertion des lignes de commandes dans la baie

En parallèle, dès la capacité des LUNs saisie, un suivi de la volumétrie apparaît. D'un coup d'œil, on peut visualiser la volumétrie allouée mensuellement ou annuellement et proposer des estimations de l'évolution au plus juste. On peut de plus réajuster le besoin en volumétrie, s'il y avait une éventuelle dérive.



Evolution mensuelle de la volumétrie à Marin

Cet outil a été utilisé quotidiennement lors de la phase de migration et sert toujours actuellement, de manière plus sporadique, durant l'exploitation.

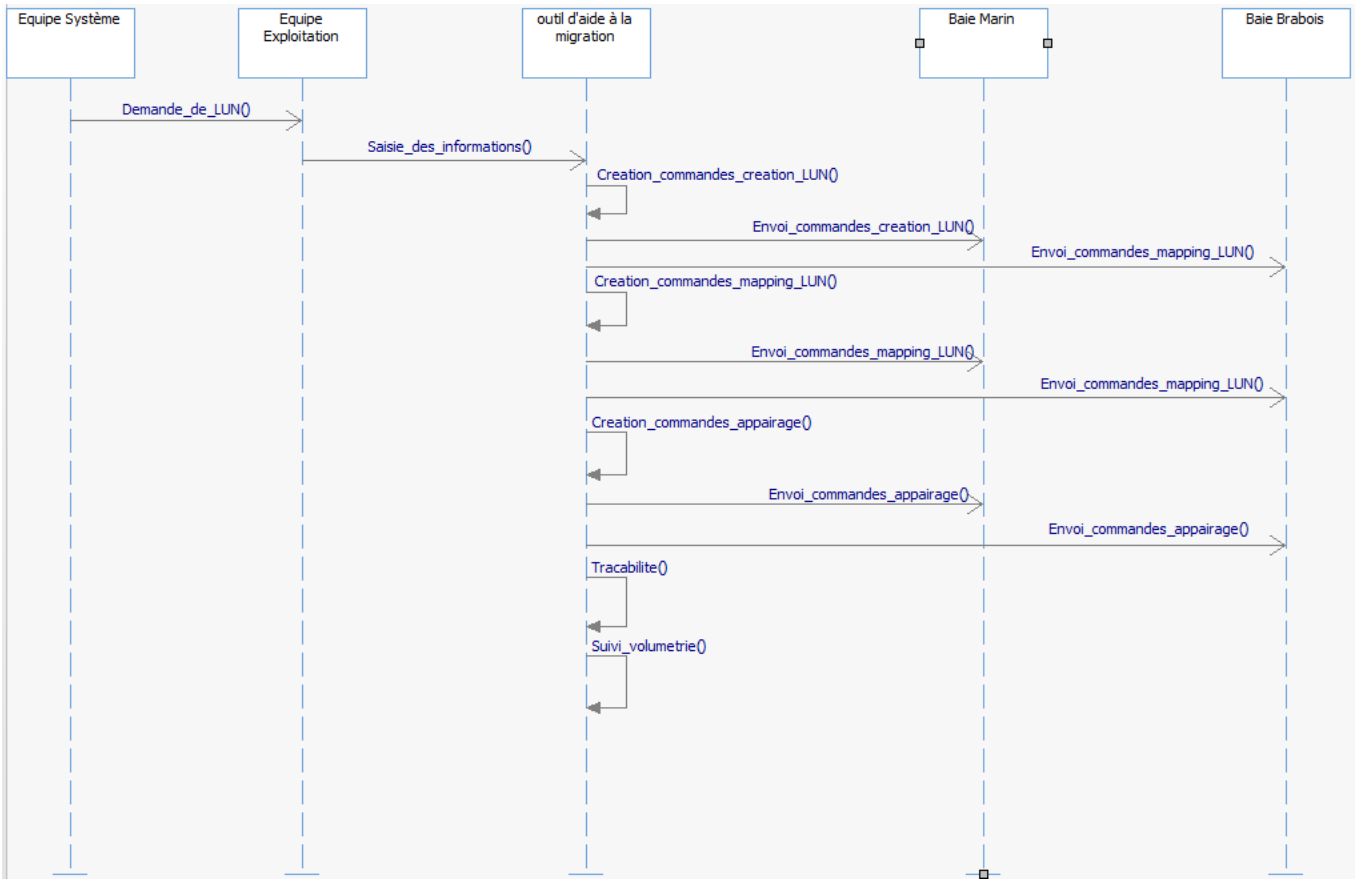
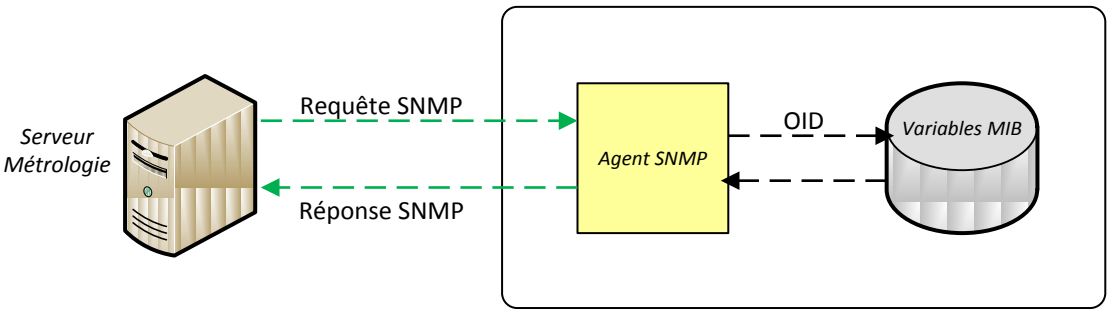


Diagramme de séquence de création d'une LUN répliquée

3.3.2 Métrologie

Afin de suivre l'évolution des performances des baies de stockage au cours de leur migration puis sur l'ensemble de leur exploitation, j'ai mis en place un serveur de métrologie.

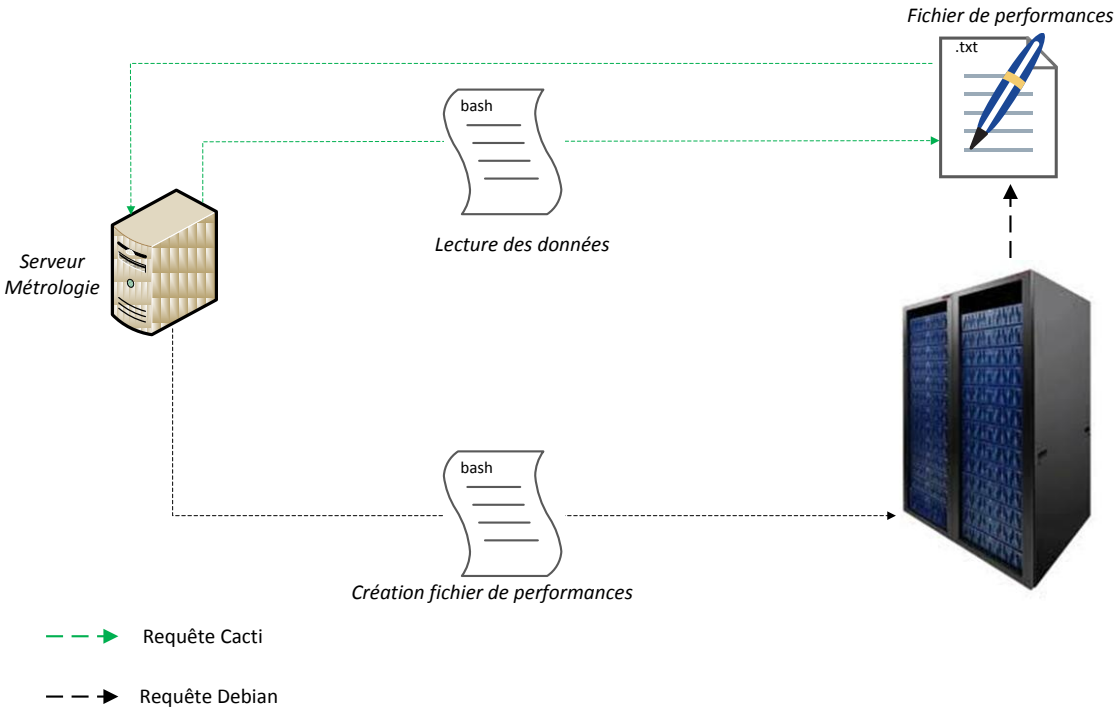
Cacti, par défaut, propose un système de polling, qui, toutes les cinq minutes, interroge l'agent SNMP de l'équipement à surveiller, en ayant au préalable précisé l'OID[24]. Ceci fait, il va pouvoir trouver la ressource adéquate dans une base de données MIB[25]. Cette technique simple d'accès – il suffit, en effet, de saisir l'OID désiré – m'a permis de mettre en place notamment la surveillance de la charge CPU[26] ou le trafic sur le NAS ou la DXi ou la température sur les switches. Cependant, la MIB du SAN étant trop pauvre, elle n'a pu être utilisée sur les baies HDS.



SNMP via Cacti

Afin de remédier à ce problème, j'ai exploité une caractéristique de la baie qui permet de créer un fichier de performances. Dans ce dernier, on trouve l'ensemble des informations qui nous intéresse, à savoir la charge CPU, le remplissage du cache, le nombre d'I/O sur les pools, LUNs ou RAID Groups. Ces données sont malheureusement « noyées » au milieu de milliers d'autres. Il est donc nécessaire de créer deux scripts. Le premier est exécuté toutes les cinq minutes par un job planifié et exige de la baie qu'elle crée le fichier de performances. Le second est exécuté par Cacti et permet de rechercher la valeur désirée dans le fichier précédemment édité. Il ne reste plus alors sous Cacti qu'à configurer un graphique adapté.

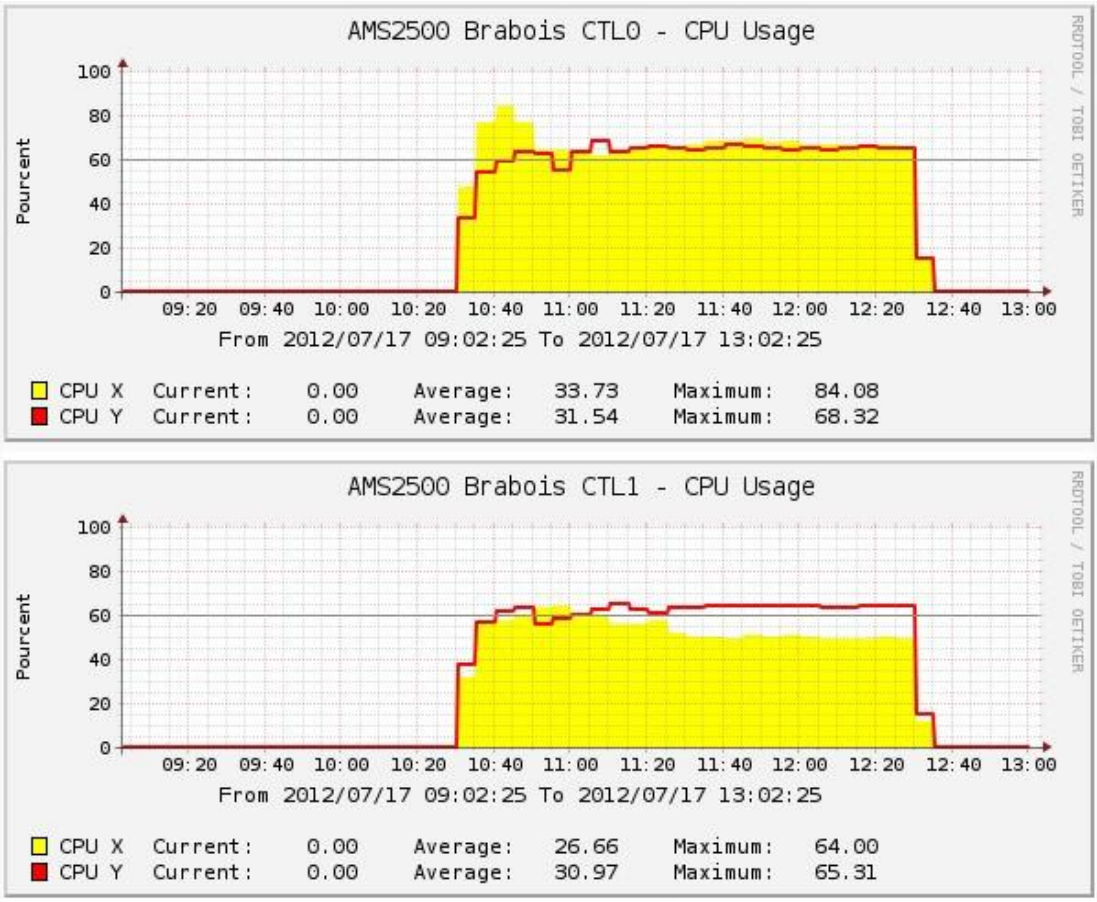
[24] Object Identifier : Identifiant unique d'une resource que l'on veut surveiller
[25] Management Information Base
[26] Central Processor Unit



Polling scripté via Cacti

On va dès lors pouvoir s'assurer que l'on respecte les préconisations du constructeur, à savoir une charge CPU < 60%, un remplissage du cache < 35% ou un taux d'écriture dans le cache à 100% (ce qui signifie qu'il n'est pas saturé). De plus, il est possible de repérer si les ressources systèmes viennent à montrer leur limite et donc envisager l'achat de mémoire vive supplémentaire par exemple. Enfin, il est indispensable de surveiller le nombre d'I/O pour superviser l'exécution de jobs sous Opcon. En effet, si on observe un pic d'I/O à une heure précise, on décidera donc de décaler les jobs afin de ne pas écrouler l'équipement.

Par exemple, on a pu observer sur la capture ci-dessous la montée en charge brutale du CPU. En effet, le 17 Juillet vers 10h30, un disque de la baie de Brabois est tombé en panne ce qui a entraîné la réécriture automatique de ses données sur un disque de *spare* (de rechange). On a donc observé un taux d'utilisation CPU légèrement supérieur aux 60% préconisé par le constructeur mais ce, pendant 2 heures seulement.



Charge CPU de la baie AMS 2500 de Brabois

Conclusion

A l'heure de conclure sur ce stage, je peux dire qu'il y a eu deux phases distinctes. Les débuts ont été compliqués car j'ai dû faire mes preuves avant qu'on ne me laisse plus d'autonomie et de possibilités d'intervenir sur le projet. Mais ensuite, j'ai pu totalement m'épanouir, jouissant d'une confiance totale de la part de l'équipe de l'Exploitation.

J'ai éprouvé quelques difficultés, surtout au début, dans l'assimilation d'un vocabulaire technique très dense. De plus, j'ai dû beaucoup utiliser les systèmes d'exploitation du monde libre – notamment pour scripter - qu'on ne nous avait jusqu'alors pas enseigné. Cela a cependant été l'occasion pour moi de mettre en avant une qualité indispensable dans un tel métier : la curiosité. En outre, le fait de travailler au sein d'un CHU exige une concentration de tous les instants compte tenu des patients qui peuvent directement être impactés par une erreur faite, ici, à la DSI.

Mon sentiment sur ce stage est au final très positif. D'un point de vue technique, j'ai pu découvrir le domaine du stockage dans une architecture très développée en participant à un projet de grande ampleur. Bien que débutant, j'ai signé un CDI dans une SSII dont la mission commence dès le lendemain de la fin de mon stage, preuve que ce projet a une réelle résonnance dans le monde professionnel.

Bibliographie

Sites Internet :

http://www.chu-nancy.fr/

http://forums.cacti.net/

Documents :

Dossier_Exploitation_Salles_Informatiques.doc

PrésentationExploitation20100505.doc

ORGANIGRAMME-DSI-tel-10012012.pdf

Architecture de sauvegarde CHU-Nancy_20120201.vsd

CHU Nancy Schémas Architecture et Cablage V00.03.vsd

Annexes

Annexe 1 : Méthodologie de création et mapping de LUNs

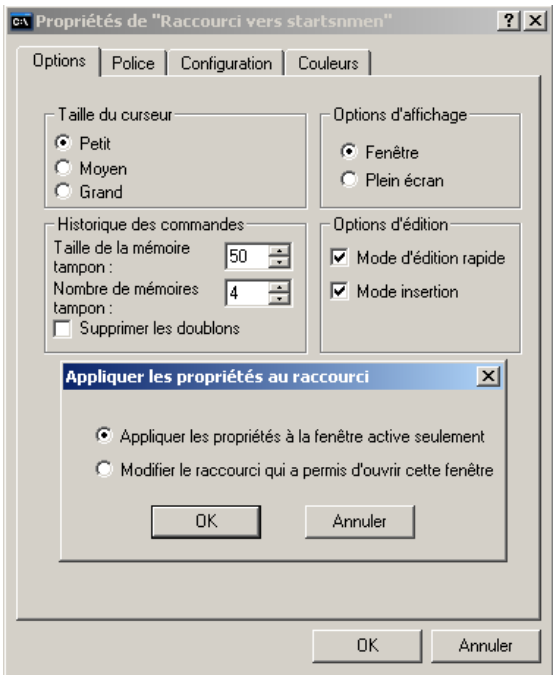
1. Introduction :

Le processus utilisé pour la création de LUNs est semi-industriel. En effet, la saisie de la documentation sous le tableur Excel permet la création automatique des commandes à copier directement dans l'éditeur. Le but est ici d'automatiser le processus de création et de mapping des LUNs afin de réduire au maximum les erreurs de saisie.

2. Démarche :

2.1 Prérequis :

Dans un premier temps, il convient d'ouvrir l'éditeur de commandes (C:\Program Files\Storage Navigator Modular 2 CLI\startsnmen.bat) et de le paramétrer. Pour cela, cliquer droit sur la barre de titres de l'éditeur puis propriétés. Cocher " Mode d'édition rapide" et valider la fenêtre pop-up. Dès lors, il est possible de copier dans l'éditeur en sélectionnant puis en appuyant sur Entrée et de coller en cliquant droit.



Ensuite, il faut ouvrir le fichier Excel permettant la création de LUNs (\\aptus3\dsi-exploitation$\Integration\Services_Operationnels\Stockage\PROJET_AMS2500\Architecture\AMS2500_Lun_v5.xlsx).

N.B. : Un lien vers CLI est disponible en D2 sur les feuilles du fichier Excel.

2.2 Création du HostGroup

Avant toute création de LUN, il faut s'assurer que le HostGroup auquel elle sera rattachée a bien été créé. Pour cela, dans l'éditeur de commandes, taper :
auhgdef -unit *Baie* -refer

```
SNM2>auhgdef -unit M_AMS -refer
Port 0A
  Group  Host Group Name
      0  GEMMA
      1  Albarine_hc
      2  Amonre
      3  Aquila
      4  Aude
      5  Bleone_hc
      6  Vios2_Gardon_hc
      7  Nonette_hc
      8  Vios1_Gardon_hc
```

S'il n'apparaît pas dans la liste, il faut le créer. Pour cela, ouvrir le fichier \\aptus3\dsi-exploitation$\Integration\Services_Operationnels\Stockage\PROJET_AMS2500\Architecture\BROCADE_configuration_v31.xlsx, feuille "Création HG". Compléter l'ensemble des champs se situant à gauche de la barre jaune, à l'exception du champ "statut création" qui permettra, une fois l'opération terminée, de surligner les lignes de commandes en vert. Il est important de noter que la création d'un HostGroup nécessite le remplissage de 2 lignes, une pour chaque Fabric. Lorsque toutes les colonnes seront dûment complétées, les commandes apparaîtront.

⚠ **La saisie des ports doit respecter cette syntaxe : "0 B" et non : "0B".**

⚠ **Par défaut, Excel interprète la saisie de "0 A" et "1 A" comme des dates (respectivement 00/01/1900 et 00/01/1900 01:00:00). Afin d'éviter cette interprétation automatique, il convient de placer une apostrophe devant : " '0 A ".**

⚠ **Il n'est pas possible de coller plusieurs lignes simultanément de la colonne "setting de l'OS de l'hostgroup". En effet, il est nécessaire après le passage d'une commande de répondre par deux "y" successifs afin de la valider.**

Dès lors, il suffit de copier successivement le contenu des colonnes "création des hostgroups", "setting de l'OS de l'hostgroup" et "ajout d'un WWN non-connu + création nickname + ajout dans le hostgroup" et de le coller dans l'éditeur de commandes.

2.3 Création des LUNs :

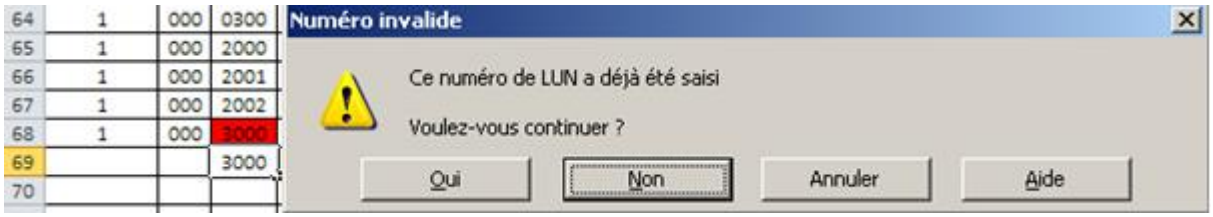
Tout d'abord, il faut choisir la baie sur laquelle on désire créer les LUNs en ouvrant la feuille appropriée (LunMarin ou LunBrabois). Ceci fait, il faut compléter de gauche à droite le tableau en commençant par la colonne POOL. Ensuite, définir un numéro de LUN unique, une capacité en MB, le port de la baie attaché au Fabric 1, celui attaché au Fabric 2 et le Hostname (et les champs Réplication dev_group et dev_name si il y a réplication). Ce n'est que lorsque ces colonnes seront dûment complétées que les commandes apparaîtront. Les autres colonnes sont facultatives dans la création des lignes de commandes.

⚠ **La saisie des ports doit respecter cette syntaxe : "0 B" et non : "0B".**

⚠ **Par défaut, Excel interprète la saisie de "0 A" et "1 A" comme des dates (respectivement 00/01/1900 et 00/01/1900 01:00:00). Afin d'éviter cette interprétation automatique, il convient de placer une apostrophe devant : " '0 A ".**

N.B. 1 : La colonne "Nombre de blocks" est calculée automatiquement à partir de la colonne "Capacity MB". De la même manière, le numéro de HLUN est défini automatiquement. Afin d'identifier plus facilement ces colonnes, elles sont grisées.

N.B. 2 : En cas de saisie d'un numéro de LUN préexistant, une fenêtre pop-up apparaît et surligne en rouge la LUN concernée.



N.B. 3 : Les baies AMS 2500 proposent l'option "Accelerated Wide Stripping Mode" permettant la répartition de la LUN sur l'ensemble des Raid Groups. Elle offre des performances accrues en multipliant les accès mais contraint la LUN à avoir une taille spécifique afin de ne pas "gaspiller" de l'espace (voir tableau ci-dessous). En effet, l'espace utilisé, option activée, est multiple de 24 Go à Marin et de 16 Go à Brabois.

Marin	Option activée	Option désactivée
Taille LUN (Go)	10	10
Espace utilisé (Go)	24	10

Dès lors, l'option pourra être désactivée sur une LUN de petite taille si celle-ci n'exige pas de hautes performances. Une telle LUN (<20 Go à Marin, <15 Go à Brabois) verra son champ " Acc. Wide Str. Mode" surligné en orange s'il est à "oui".

POO	LUN	Capacity MB	Nombre de blocks	Port 1	Port 2	Host Groups	HLUN	hdisk	VG	Description LUN	Point de montage ou serveurs ESX	Réplication dev_group	Réplication dev_name	Acc. Wide Str. Mc
000	0031	10 000	20 480 000	0 C	1 C	Moselle	0003	hdisk11	vgr22			gipprod	P0062S0031	oui

Ainsi, dès la saisie d'une taille de LUN, le champ "Taille réelle (MB)" affiche la taille réellement consommée sur la baie.

N.B. 4 : Lors de la création d'une LUN répliquée, la perte d'espace est nulle si et seulement si la taille de la LUN est à la fois multiple de 24 et de 16 Go. Il est donc rarement possible d'éviter totalement ce gaspillage. Toutefois, le champ "Taille optimisée (MB)" calcule automatiquement (dès lors que le champ "Acc. Wide Str. Mode" est à "oui" et que "Réplication dev_name" est non-vide) la taille de LUN afin de minimiser au maximum la perte d'espace.

La précédente saisie permet la création automatique des commandes nécessaires à la création de LUNs et à leur mapping. Dès lors, il suffit de copier successivement le contenu des colonnes "Création LUN", "Mapping Lun Port 1" et "Mapping LUN Port 2" et de le coller dans l'éditeur de commandes. La colonne "horcm 1x" sera utile en cas de réplication (cf. 2.5).

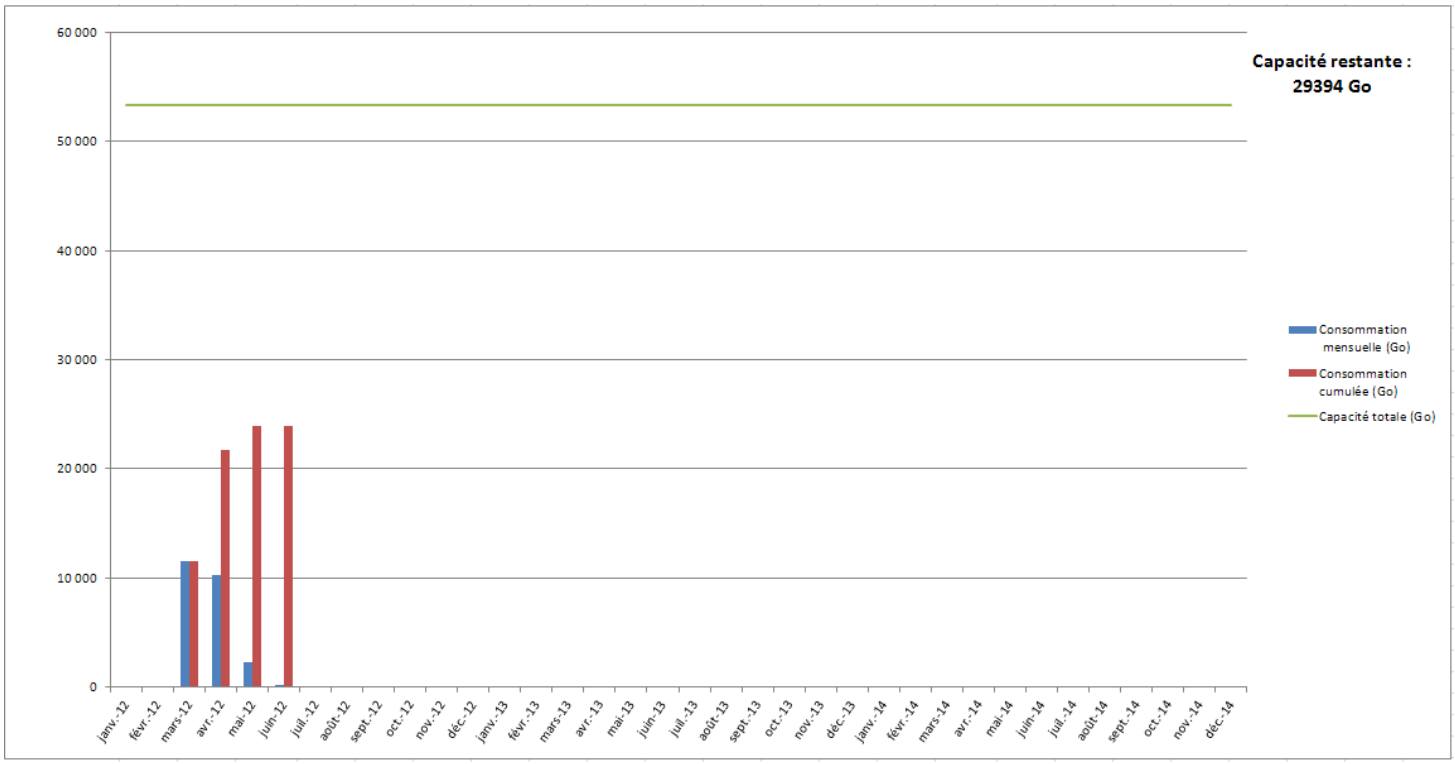
N.B. : Alors que l'interface graphique permet de mapper une LUN sur 2 ports à la fois, en mode commande, il est nécessaire de passer une commande pour chacun des ports.

```
SNM2>Echo y | auluadd -unit B_AMS -lu 21 -dppoolno 0 -size 40000m -widestriping enable -fullcapacity enable
Are you sure you want to set the logical unit? (y/n [n]): The logical unit has been set successfully.

SNM2>Echo y | auhgmap -unit B_AMS -add 0 C -gname Tarn -hlu 0 -lu 21
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.

SNM2>Echo y | auhgmap -unit B_AMS -add 1 C -gname Tarn -hlu 0 -lu 21
Are you sure you want to add the mapping information? (y/n [n]): The mapping information has been set successfully.
```

Un copier-coller d'un trop grand nombre de commandes d'un seul coup peut entraîner une perte d'informations due à la taille limitée du presse-papiers. Il est conseillé de ne pas passer plus d'une dizaine de LUNs à la fois.

Une fois les commandes passées, il convient de passer le "Statut création" à 1 surlignant les lignes de commandes en vert indiquant qu'elles ont bien été passées.

Une fois toutes ces informations saisies, il est possible de suivre l'évolution de la consommation en termes de volumétrie dans les onglets "Cons. Mens." et "Cons. Ann.". Le premier rend compte de la consommation chaque mois jusqu'en 12/2014 ainsi que la consommation cumulée et le second applique le même principe pour les années jusqu'en 2016. Sur chacun d'eux, il est en outre possible de visualiser la capacité restante.



2.4 Validation :

Afin de valider les commandes saisies, il suffit d'entrer celle-ci :
auluref -unit *Baie –g/m/t*

```
SNM2>auluref -unit B_AMS -m
                              Stripe  RAID     DP    RAID                          Number
   LU       Capacity          Size    Group    Pool  Level        Type             of Paths  Status
   21       40000.0 MB         256KB    N/A       0    5( 7D+1P)  SAS                     2  Normal
```

Grâce à cette commande, on peut vérifier que la taille demandée pour la LUN a bien été allouée et sur le bon Pool. Il est également possible de valider le mapping avec cette commande :
auhgmap –unit *Baie* –refer

```
SNM2>auhgmap -unit B_AMS -refer
Mapping Mode = ON
Port  Group                                  H-LUN    LUN
  0C  005:Tarn                                   0     21
  1C  005:Tarn                                   0     21
```

On peut alors visualiser que la LUN précédemment créée est affectée aux bons ports et Hostgroup.

2.5 Création des fichiers horcm :

Dans le cas où les LUNs créées seraient répliquées, les informations saisies dans le fichier *AMS2500_Lun_v5.xlsx* permettent de créer les commandes à entrer dans les fichiers horcm (respectivement horcm10 dans la feuille LunMarin et horcm11 dans la feuille LunBrabois).

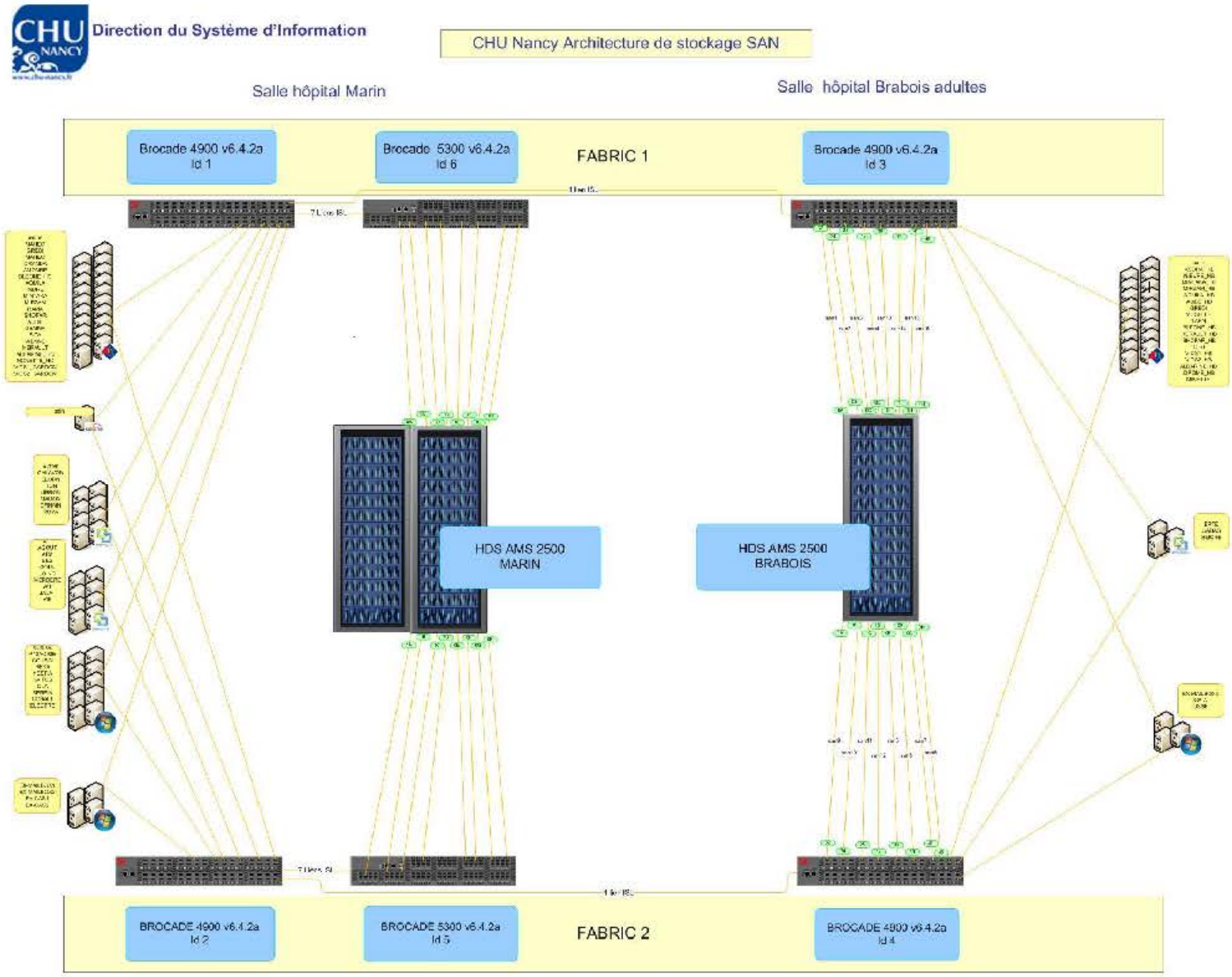
Avant toute modification des fichiers, il est nécessaire de créer une backup. Pour cela, dans Gemma et Gredi (\\aptus3\dsi-exploitation$\Integration\Services_Operationnels\Stockage\PROJET_AMS2500\CCI), créer un dossier du type "aaaammjj" et y glisser une copie de *horcm10.conf* et *horcm11.conf*. Ensuite, il suffit de les ouvrir grâce à Notepad++ et de copier le contenu de la colonne "horcm 1x" (dev_group, dev_name, port#, Target ID et LU#) et de le coller dans la partie HORCM_DEV en respectant l'ordre alphabétique établi.

```
HORCM_DEV
#/=============================================================================/
#dev_group        dev_name          port#          TargetID        LU#        MU#
#/-----------------------------------------------------------------------------/
#/ Albarine_hc ==> Albarine_hb ------------------------------------------------/
glimsprod         P0173S0112        CL1-A             0           0173
glimsprod         P0174S0113        CL1-A             0           0174
glimsprod         P0175S0114        CL1-A             0           0175
glimsprod         P0176S0115        CL1-A             0           0176
glimsprod         P0177S0116        CL1-A             0           0177
glimsprod         P0178S0117        CL1-A             0           0178
glimsprod         P0179S0118        CL1-A             0           0179
#/ Amonre ==> Saule -----------------------------------------------------------/
bnpcprd           P0073S0038        CL1-A             0           0073
capprd            P0074S0039        CL1-A             0           0074
```

N.B. : Il reste nécessaire de remplir manuellement la partie HORCM_INST avec les informations dev_group, ip_address et service.

```
HORCM_INST
#/=============================================================================/
#dev_group         ip_address              service
#/-----------------------------------------------------------------------------/
anapath                                   horcm11
aquila                                    horcm11
aude                                      horcm11
bmac                                      horcm11
bnpcprd                                   horcm11
boxi                                      horcm11
capprd                                    horcm11
```

Annexe 2 : Architecture de stockage SAN du CHU



Annexe 3 : Script de relevé des performances des RAID Groups

```
#!/bin/bash

date_act=$(date +%s)
date_fich0=$(date -r $3"/pfm00000.txt" '+%s')
date_fich1=$(date -r $3"/pfm00001.txt" '+%s')
ecart0=$(expr $date_act - $date_fich0)
ecart1=$(expr $date_act - $date_fich1)

if [ $ecart0 -lt $ecart1 ] && [ $ecart0 -ge 60 ]; then
fichier="pfm00000.txt"
elif [ $ecart0 -lt $ecart1 ] && [ $ecart0 -lt 60 ]; then
fichier="pfm00001.txt"
elif [ $ecart1 -lt $ecart0 ] && [ $ecart1 -ge 60 ]; then
fichier="pfm00001.txt"
elif [ $ecart1 -lt $ecart0 ] && [ $ecart1 -lt 60 ]; then
fichier="pfm00000.txt"
else
fichier="pfm00000.txt"
fi

awk -f "/usr/share/cacti/site/scripts/RG_info.awk"  -v RG=$1 ctl=$2 $3/$fichier
```

```
BEGIN{ORS=" "
}
{ if (/---- RG Information ----/){
    for(NR;$3!="Read";NR++){
        if(($2==RG)&&($1==ctl)){
        print "IOPS_RG:"$3" R_Hit_RG:"$6" W_Hit_RG:"$7;}
        getline}
    }
}
END{
}
```

RESUME :

Le stage consiste à la mise en œuvre du nouveau système de stockage SAN du CHU de NANCY.

Il s'agit de mettre en place de nouvelles baies sur les sites de production et de secours, puis de migrer les données des anciennes baies vers les nouvelles baies tout en maintenant en condition opérationnelle les systèmes informatiques.

MOTS-CLES :

Stockage, SAN, migration, système d'information

ABSTRACT :

The internship is about the implementation of the new storage system in "CHU de Nancy".

The purpose is to set up new storage arrays in two places: production site and backup site. Then, we have to migrate the data from the old arrays to the new ones and keep the information system operational.

KEYWORDS :

Storage, SAN, migration, information system