Sécurisation de l'environnement matériel d'exécution d'un dossier patient informatisé

Guillaume Marin

CONSERVATOIRE NATIONAL DES ARTS ET METIERS

CENTRE REGIONAL DE POITOU CHARENTES

_____

MEMOIRE

En vue d'obtenir

Le DIPLOME d'INGENIEUR CNAM

SPECIALITE : Informatique

OPTION : Système d'Information

par

Guillaume MARIN

_____

« Sécurisation de l'environnement matériel d'exécution d'un dossier patient informatisé »

Soutenu le 13 décembre 2013.

_____

Jury

PRESIDENT : Monsieur Stéphane Natkin, Professeur CNAM Paris

MEMBRES : Madame Marie-Christine Lafaye, Responsable filière Informatique CNAM Poitou-Charentes

Monsieur Mourad Rabah, Maître de Conférences, IUT La Rochelle

Madame Joëlle Gabilleau, Directrice du Centre Hospitalier du Blanc

CONSERVATOIRE NATIONAL DES ARTS ET METIERS

CENTRE REGIONAL DE POITOU CHARENTES

_____

MEMOIRE

En vue d'obtenir

Le DIPLOME d'INGENIEUR CNAM

SPECIALITE : Informatique

OPTION : Système d'Information

par

Guillaume MARIN

_____

« Sécurisation de l'environnement matériel d'exécution d'un dossier patient informatisé »

Soutenu le 13 décembre 2013.

_____

Jury

PRESIDENT : Monsieur Stéphane Natkin, Professeur CNAM Paris

MEMBRES : Madame Marie-Christine Lafaye, Responsable filière Informatique CNAM Poitou-Charentes

Monsieur Mourad Rabah, Maître de Conférences, IUT La Rochelle

Madame Joëlle Gabilleau, Directrice du Centre Hospitalier du Blanc

REMERCIEMENTS

Ce mémoire est l'aboutissement d'un long parcours de formation au sein du CNAM, ainsi que la concrétisation d'une expérience professionnelle développée en grande partie au Centre Hospitalier du Blanc. Je tiens donc à remercier tout particulièrement les différents acteurs du CNAM, non pas uniquement parce que c'est l'usage, mais parce que cette grande institution nous donne la possibilité à tous, de poursuivre des études supérieures de grande qualité, parallèlement à la vie active.

Je tiens aussi à remercier vivement les directeurs successifs des structures dans lesquelles j'ai évolué, et notamment les directeurs du Centre Hospitalier, pour m'avoir soutenu et accordé leur confiance à l'occasion de la réalisation de nombreux projets dont celui m'ayant permis la réalisation de ce mémoire.

Je remercie aussi Monsieur le Président du Jury, Monsieur Stéphane Natkin, ainsi que tous les membres qui le composent et qui m'ont fait l'honneur d'être présents et d'avoir consacré du temps à ce projet.

Je remercie également Monsieur Mourad Rabah d'avoir assuré avec bienveillance le tutorat de ce mémoire.

Finalement, ce travail est un projet personnel qui a entrainé tout mon entourage et tout particulièrement mon épouse tout au long de ces années. Je la remercie donc pour sa patience et son soutien sans lesquels ce formidable mais long projet n'aurait pu être mené à bien.

LISTE DES ABREVIATIONS

ADSL : Asymetric Digital Suscriber Line

AES : Advanced Encryption Standard

BC : Business Continuity

CCTP : Cahier des Clauses Techniques Particulières

CD-ROM : Compact Disc - Read Only Memory

CDI : Centre de Documentation et d'Information

CDP : Continuous Data Protection

CH / CHU : Centre Hospitalier / Centre Hospitalier Universitaire

CHT : Communauté Hospitalière de Territoire

CPU : Central Processing Unit (Microprocesseur principal d'un ordinateur)

DAT : Digital Audio Tape

DECT : Digital Enhanced Cordless Telecommunications

DHCP : Dynamic Host Configuration Protocol

DIM : Département d'Information Médicale

DMP : Dossier Médical Personnel

DMS : Durée Moyenne de Séjour

DPI : Dossier Patient Informatisé

DR : Disaster Recovery

DVD-ROM: Digital Versatile Disc - Read Only Memory

EAI : Entreprise Application Integration (Intégration d'Applications d'Entreprise)

EDI : Electronic Data Interchange / Echange de Données Informatisé

EHPAD : Etablissement Hébergeant des Personnes Agées Dépendantes

FC : Fibre Channel

FCoE : Fibre Channel Over Ethernet

FOM : FailOver Manager

GAP : Gestion Administrative du Patient

GSM : Global System for Mobile communication

HA : High Availability

HAD : Hospitalisation A Domicile

HP : Hewlett-Packard

ICA : Independent Computing Architecture

IFSI : Institut de Formation en Soins Infirmiers

IFAS : Institut de Formation d'Aides Soignants

iSCSI : Internet Small Computer System Interface

LAN : Local Area Network

LUN : Logical Unit Number

MAC : Media Access Control

MCO : Médecine, Chirurgie, Obstétrique

MOA : Maîtrise D'Ouvrage

MOE : Maîtrise d'OEuvre

MDT : Mean Down Time

MTD : Maximum Tolerable DownTime

MTTRep : Mean Time To Repair

MTTRes : Mean Time To Restore ou Mean Time to Recover

MUT : Mean Up Time

NAS : Network Attached Storage

NFS : Network File System

NUS : Network Unified Storage

OEM : Original Equipment Manufacturer

PC : Personal Computer

PeSIT : Protocole d'Echanges pour un Système Interbancaire de Télécompensation

PESv2 : Protocole d'Echange Standard Version 2

PGI : Progiciel de Gestion Intégrée

PRA : Plan de Reprise d'activité

PtoV / P2V : Physical-to-Virtual

QoS : Quality of Service

RAID : Redundant Array of Inexpensive Disks

RAM : Random Access Memory

RDP : Remote Desktop Protocol

RIS : Radiology Information System

RPO : Recovery Point Objective

RTO : Recovery Time Objective

SAN : Storage Area Network

SAS : Serial Attached SCSI

SCSI : Small Computer System Interface

SDN : Software Defined Network

SGBD : Système de Gestion de Base de Données

SI / SIH : Système d'Information / Système d'Information Hospitalier

SMB : Server Message Block

SMUR : Service Mobile d'Urgence et de Réanimation

SPOF : Single Point Of Failure

SR : Sous Répartiteur

SSIAD : Service Soins Infirmiers A Domicile

SSR : Soins de Suite et de Réadaptation

TCO : Total Cost of Ownership

TCP/IP : Transmission Control Protocol / Internet Protocol

UE : Unité d'Enseignement

USB : Universal Serial Bus

USC : Unité de Surveillance Continue

USLD : Unité de Soins de Longue Durée

VDI : Virtual Desktop Initiative

VLAN : Virtual Local Area Network (ou Virtual LAN)

VM : Virtual Machine

VPN : Virtual Private Network

VSAN : Virtual Storage Area Network (ou Virtual SAN)

VTL : Virtual Tape Library

VtoV / V2V : Virtual-to-Virtual

XVA : XenServer Virtual Appliance

WIFI : WIreless FIdelity

WRT : Work Recovery Time

ZHTCD : Zone d'Hospitalisation de Très Courte Durée

SOMMAIRE

INTRODUCTION

En 2008, le Centre Hospitalier du Blanc a inscrit comme objectif prioritaire pour l'établissement, l'informatisation du dossier du patient. Cela devait notamment couvrir la mise en place de tout le suivi du circuit du médicament pour lequel l'établissement s'était engagé dans le cadre du « contrat de bon usage du médicament » [1] mais, plus globalement, tous les aspects liés à la prise en charge médicale et paramédicale du patient : « Un dossier médical est constitué pour chaque patient hospitalisé dans un établissement de santé public ou privé. Ce dossier contient au moins […] les informations formalisées recueillies lors des consultations externes dispensées dans l'établissement, lors de l'accueil au service des urgences ou au moment de l'admission et au cours du séjour hospitalier » [2].

Cet ensemble d'applications allait donc devenir critique pour la prise en charge du patient hospitalisé. Il est alors apparu comme nécessaire de mettre en place une infrastructure permettant d'assurer une disponibilité maximum compatible avec le progiciel qui allait être choisi, tout en se conformant aux possibilités financières et humaines d'un Centre Hospitalier de taille modeste. Il ne s'agit pas ici de la mise en place d'un Plan de Reprise d'Activité (PRA) mais de la mise en place d'une des briques techniques qui servira ensuite de socle indispensable à sa mise en place. Cette brique technique aura consisté en la virtualisation d'un ensemble de serveurs et du stockage associé sur une plateforme en haute disponibilité.

Le projet présenté au travers de ce mémoire s'inscrit dans un processus global et coordonné, composé d'un ensemble de mises à niveau matérielles et logicielles cohérentes, nécessaires à l'atteinte des objectifs, que nous aborderons au cours de la première partie de ce document consacrée à l'environnement de développement du projet.

Nous passerons ensuite en revue l'état de l'art des solutions de virtualisation et de haute disponibilité notamment des serveurs et du stockage associé, avant d'en venir à la mise en application qui a pu en être faite dans le cadre de la mission qui m'a été confiée au CH du Blanc.

1. CONTEXTE DU PROJET

L'objectif premier et moteur de bon nombre d'évolutions technologiques pour le Centre Hospitalier du Blanc aura été la mise en place du dossier patient informatisé. Projet qui nécessitait une disponibilité maximum de l'information médicale et paramédicale du patient, tout au long de son séjour.

1.1 PRESENTATION DU CENTRE HOSPITALIER DU BLANC

« Le plus vieux texte faisant mention d'une maison hospitalière au Blanc, date de 1524 » [3] mais c'est plus récemment, en 1713, que sont posées les premières pierres de l'actuel hôpital. C'est une structure à « taille humaine » située dans une zone peu peuplée (29,8 habitants au km² dans le canton de Le Blanc et 19,2 habitants au km² si on inclut les 4 cantons limitrophes les plus proches [4] de la ville de Le Blanc) et éloignée des autres structures hospitalières (35 minutes du CH de Montmorillon, une heure des CH de Châteauroux et Châtellerault ainsi que du CHU de Poitiers et deux heures des CHU de Limoges et Tours). C'est probablement cet isolement géographique qui est à l'origine de la pluralité de ses activités et qui lui confère le statut de « Centre Hospitalier ». Comme nous le précise le Code de la Santé Publique, il est classé dans « les établissements publics de santé qui ne figurent ni sur la liste des centres hospitaliers régionaux, ni sur les listes d'hôpitaux locaux » [5] donc les établissements de « taille moyenne ».

Illustration 1. Extrait de la cartographie de la densité de population par canton (1999) [4]

1.1.1 ROLES ET ACTIVITES DU CH DU BLANC

Le CH du Blanc assure deux grands types de mission. Comme le précise la loi du 31 juillet 1991, « l'hôpital assure la protection sanitaire de la population du secteur qu'il dessert » [3] et, « conformément à la loi du 30 juin 1975, il héberge les personnes âgées qui ne veulent ou ne peuvent demeurer chez elles » [3]. Contribuent à ces actions un peu plus de 400 agents.

Pour remplir ces missions, le CH met à disposition de la population, des services d'hospitalisation, d'hébergement pour personnes âgées dépendantes, de consultations spécialisées ainsi qu'un service d'urgence / SMUR, un plateau technique complet ou encore un IFSI / IFAS (communément appelé Ecole d'infirmières et d'Aides Soignantes).

1.1.2 ORGANISATION DE L'ACTIVITE MEDICALE DU CH

Conformément à ses obligations, l'activité médicale se décompose en « pôles » puis, à l'intérieur de ceux-ci, en « Services » :

1) Pôle à orientation Médicale composé de 3 services :

- 33 lits de Médecine (« Le service est spécialisé notamment en cardiologie (doppler, électrocardiographie, holter, échographie cardiaque) et en gastro-entérologie (proctologie, échographie abdominale, endoscopie digestive, traitement du diabète » [3]) ainsi que 3 lits d'Unité de Soins Continus rattachés au service de Médecine.

- 30 lits de Soins de Suite et Réadaptation.

- Un service d'imagerie médicale (radiographie, échographie, scanner).

2) Pôle à orientation Chirurgicale composé de 4 services :

- 15 lits de Chirurgie complète et ambulatoire (« Chirurgie générale, viscérale, orthopédique, plastique et reconstructive » [3] où la chirurgie pratiquée est polyvalente et où y sont traités « tous les patients en observation nécessitant des explorations fonctionnelles, des soins post-opératoires, ophtalmologiques, etc. » [3]),
- 13 lits de Gynécologie-Obstétrique,
- Service d'urgence / SMUR auquel est rattachée une Zone d'Hospitalisation de Très Courte Durée (ZHTCD).
- Service de pharmacie.

3) Pôle Médico-social :

- 30 lits en Unité de Soins de Longue Durée (USLD),
- 145 lits d'Etablissement Hébergeant des Personnes Agées Dépendantes (EHPAD),
- 41 places réparties entre l'Hospitalisation A Domicile (HAD) et le Service Soins Infirmiers A Domicile (SSIAD).

1.1.3 LES DIFFERENTES STRUCTURES

Cette activité est répartie sur trois sites géographiques différents :

- Le Centre Hospitalier regroupe l'ensemble de l'activité dite de

Illustration 2. Vue du CH du Blanc [6]

« court séjour » (les hospitalisations) et celle dite de « moyen séjour » (les hospitalisations en Soins de Suite ou de Réadaptation) mais aussi le service d'urgence / SMUR, les consultations spécialisées, l'HAD, le SSIAD, l'IFSI, etc. Se trouve également sur ce même site, l'EHPAD « Maison de Retraite Saint-Lazare ».

- L'EHPAD « La Cubissole », également localisé à Le Blanc propose en plus de l'hébergement de personnes âgées, une Unité de Soins de Longue Durée (USLD). C'est la plus importante structure d'hébergement du Centre Hospitalier.

Illustration 3. Vue du site de la Cubissole [6]

- L'EHPAD « Résidence de l'Anglin », localisé à Concremiers (7 km de Le Blanc), assure, quant à lui, uniquement l'hébergement de personnes âgées.

Illustration 4. Vue de la Résidence de l'Anglin [6]

1.1.4 LE CENTRE HOSPITALIER DANS SON ENVIRONNEMENT

Afin de pouvoir disposer des meilleures compétences en nombre suffisant malgré une région peu attractive, une pénurie de personnels, en particulier dans certaines compétences (obstétrique, anesthésie, psychologie, etc.), et une activité qui n'est pas toujours des plus complexes, le Centre Hospitalier du Blanc a su construire des partenariats avec les établissements géographiquement les plus proches. Et, cela, au-delà

de la région administrative à laquelle il se trouve rattaché. C'est aujourd'hui le cas avec le CH de Châteauroux pour les services : SMUR, HAD ou encore le service de Gynécologie-Obstétrique mais aussi avec le CHU de Poitiers pour la Chirurgie.

Sur la fin du projet, il était alors évoqué la mise en place d'une Communauté Hospitalière de Territoire (CHT) dont l'objectif principal était la mise en place d'une coopération forte entre les établissements membres (mise en commun de ressources d'encadrement ou encore téléimagerie). Etait aussi évoqué un partenariat avec le CHU de Poitiers en vue d'une coopération pour assurer les analyses biologiques du CH du Blanc qui, jusque-là, étaient externalisées auprès d'un prestataire privé.

Concernant l'informatique, au-delà d'une éventuelle mutualisation de ressources évoquée dans le cadre de la CHT, de manière moins formelle, les responsables informatiques de l'ensemble des centres hospitaliers de la région se retrouvent périodiquement et partagent leurs connaissances et expériences au sein d'un collège régional lui-même représenté au sein d'un collège national. Depuis ce dernier, sont diffusées bon nombre d'informations stratégiques en lien notamment avec les évolutions réglementaires ou les projets nationaux du moment.

1.1.5 L'HOPITAL EN QUELQUES CHIFFRES

Après avoir dressé succinctement le panel des activités du CH du Blanc, il apparaît nécessaire d'apporter quelques brefs éclairages quand à la mesure de ses activités, aux budgets associés mais aussi à la part allouée au système d'information hospitalier.

1.1.5.1 Budget du Centre Hospitalier par grands domaines d'activité

En 2010, le budget du CH du Blanc dépassait les 29 millions d'euros ainsi répartis (par grands secteurs d'activité) :

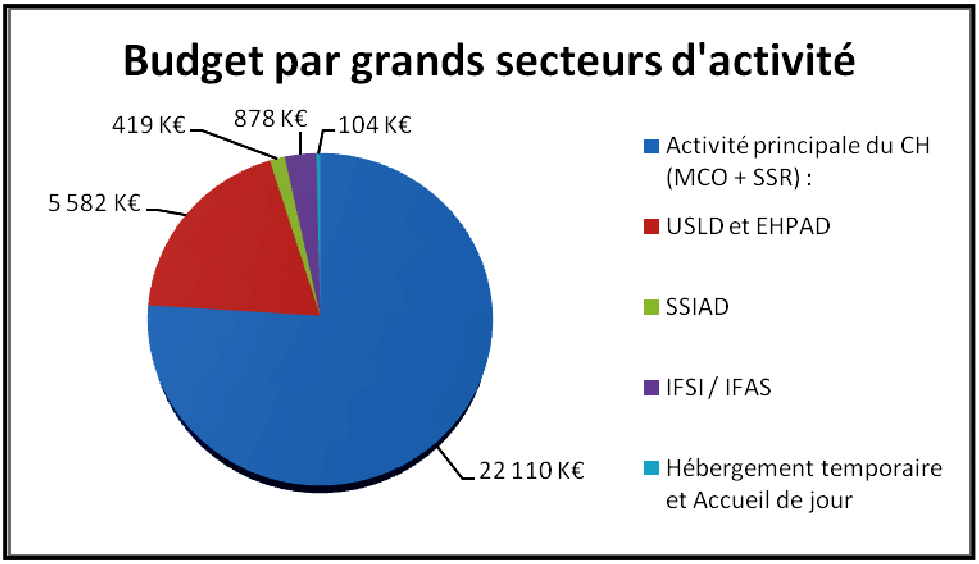


Illustration 5. Représentation du budget du CH du Blanc, issue des données du rapport d'activité [7].

Comme nous le rappellent les médias, les établissements publics de santé doivent faire face à d'importantes difficultés financières. Le CH du Blanc ne fait pas exception et présente depuis plusieurs années un exercice déficitaire.

1.1.5.2 Evolution de l'activité

Le graphique suivant, issu du « rapport d'activité et de gestion 2010 » [7], nous présente

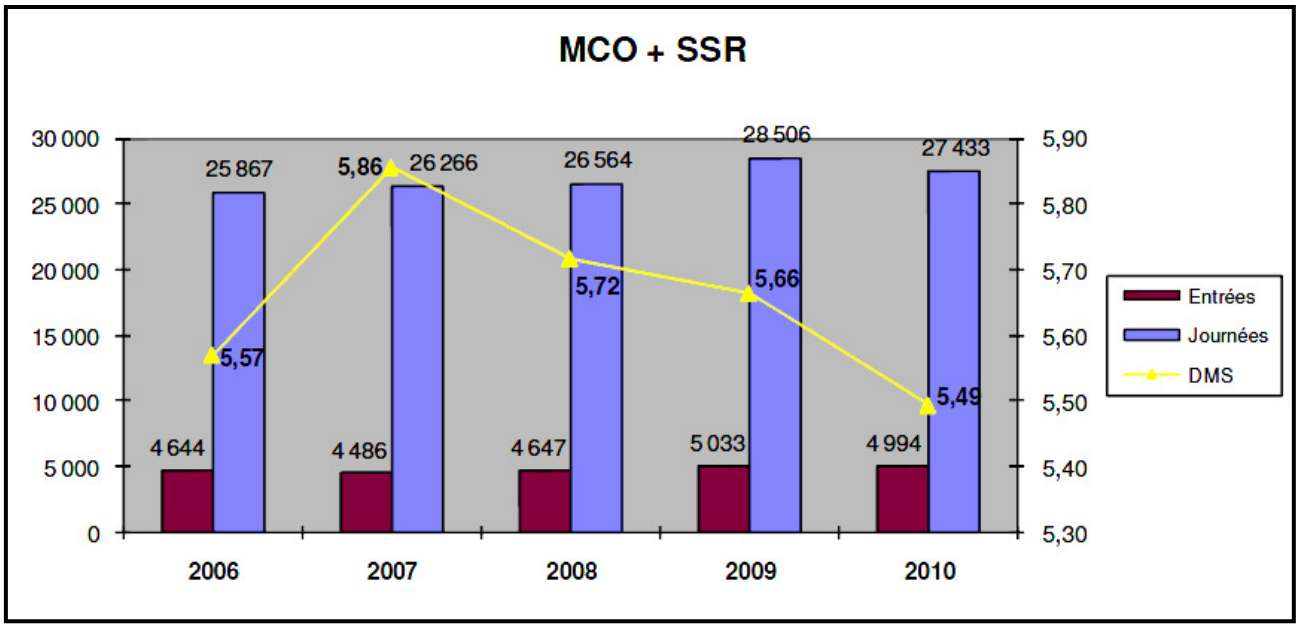


Illustration 6. Synthèse de l'activité d'hospitalisation (MCO et SSR) du CH du Blanc en nombre d'entrées / nombre de journées ainsi que la Durée Moyenne de Séjour [7]

l'évolution de l'activité d'hospitalisation des dernières années en nombre d'entrées et en nombre de journées réalisées. Il laisse apparaître un très léger accroissement de l'activité d'hospitalisation.

La progression des consultations spécialisées dites « publiques » suit sensiblement la même courbe [7] alors que le taux d'occupation des lits en secteur médico-social atteint 98,5 % [7] (toujours en 2010).

1.1.5.3 Evolution des dépenses liées au SIH

La mise en place du Dossier Patient Informatisé ainsi que l'effort porté sur l'ensemble des investissements nécessaires à la mise en place de l'infrastructure dont celle qui est l'objet de ce mémoire transparaissent au travers du graphique suivant.

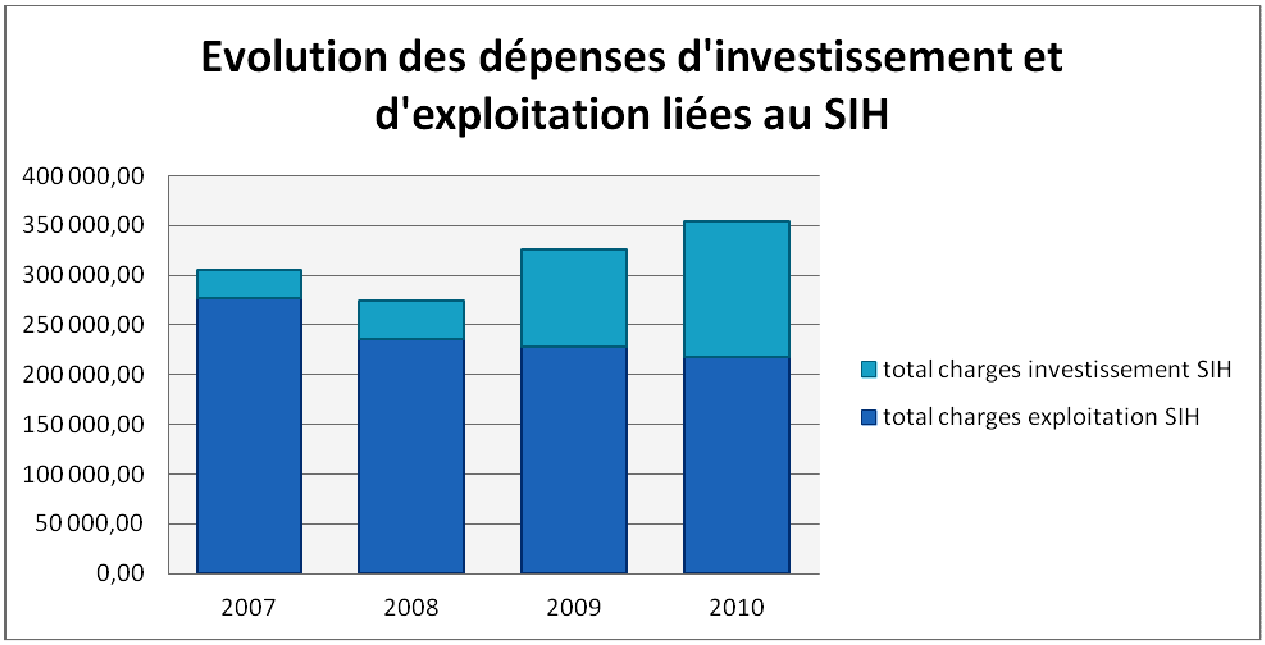


Illustration 7. Représentation de l'évolution des dépenses SIH, issue des données du rapport d'activité [7]

Nous pouvons en effet constater qu'un effort d'investissement important a été fait dès 2007 pour stabiliser la partie administrative du dossier patient puis les années suivantes pour, d'une part préparer l'architecture en vue du dossier patient puis le dossier patient lui-

même. Parallèlement, on peut remarquer une diminution importante des charges d'exploitation, fruit de la politique de rationalisation que j'ai pu intégrer à chaque projet et ce, malgré que ces chiffres intègrent d'une part la maintenance des nouvelles infrastructures (auparavant non maintenues), les charges fixes de personnels mais, d'autre part, l'équipement croissant en informatique médicale gérée, pour sa part, par le service bio-médical.

Malgré la nécessité de réaliser d'importants travaux, l'établissement a consacré en 2010 près de 8% de ses investissements au système d'information.

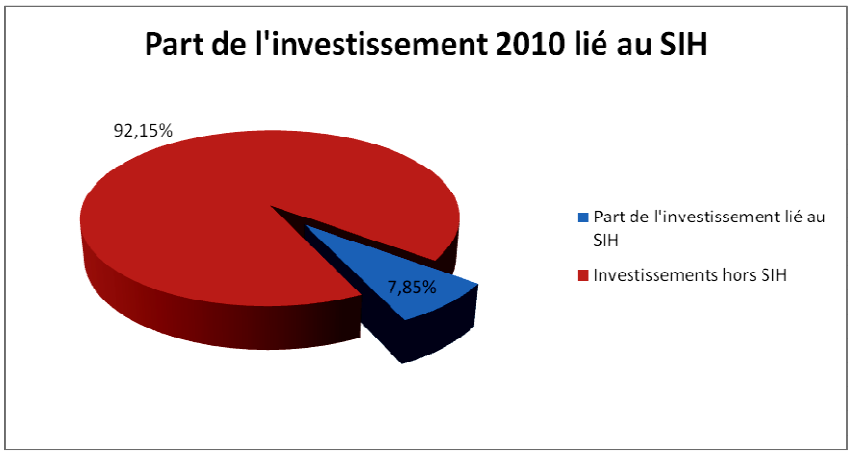


Illustration 8. Représentation de la part de l'investissement au CH du Blanc dédiée au SIH en 2010 (issue du rapport d'activité [7])

L'Observatoire des Systèmes d'Information de Santé [8], met en exergue l'effort d'investissement de l'établissement dès 2009 : celui-ci était suppérieur de 31,16% à la moyenne nationale des 213 établissements de santé de taille similaire.

1.1.6 ORGANIGRAMME DU CH DU BLANC

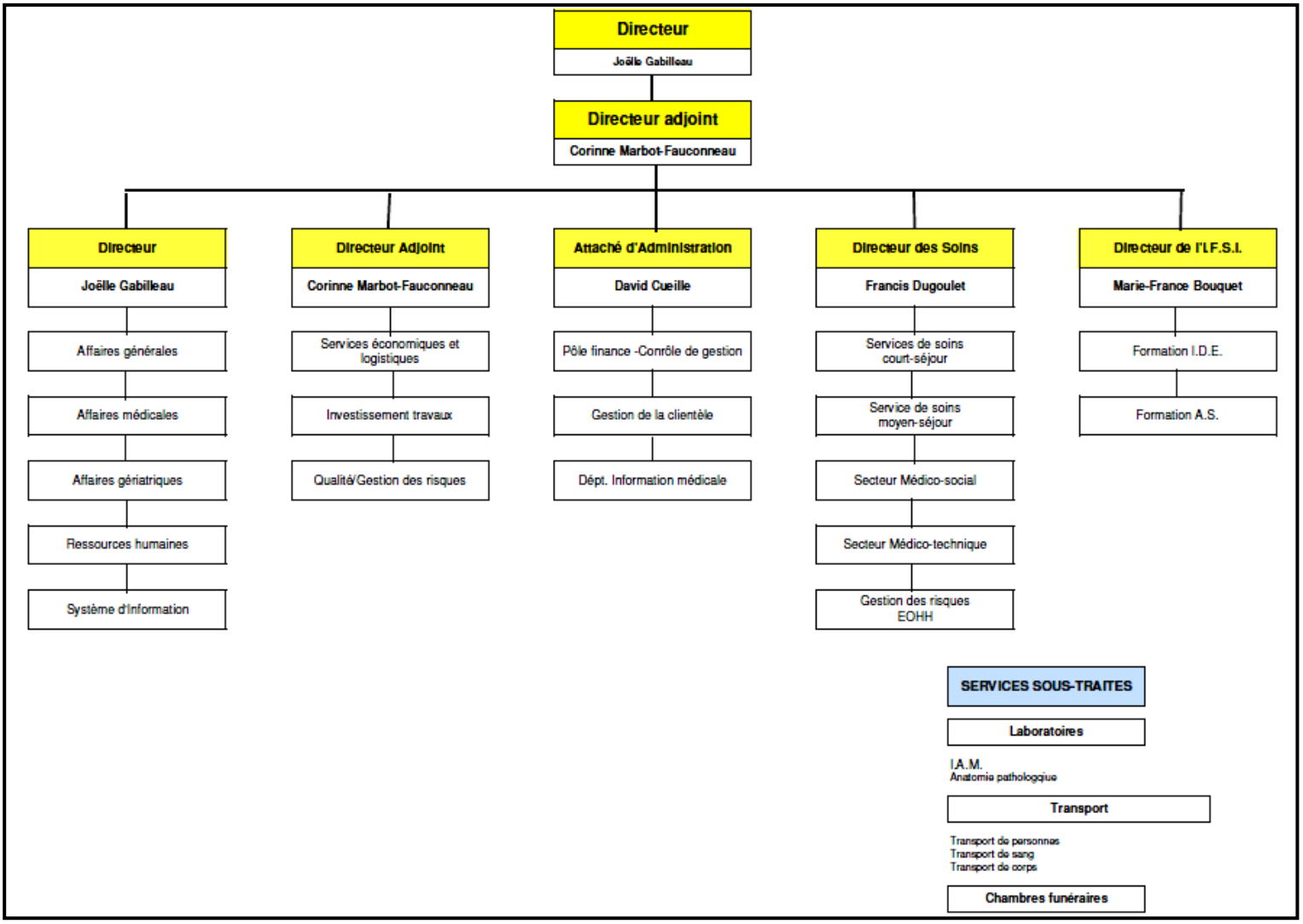


Illustration 9. Organigramme du CH du Blanc [6]

1.2 L'INFORMATIQUE AU SEIN DU CH DU BLANC.

1.2.1 L'ORGANISATION DU SERVICE INFORMATIQUE

La « direction du système d'information », est directement rattachée à la directrice du Centre Hospitalier. Je suis le responsable de ce service qui est composé de deux personnes et dont le panel de ses missions est détaillé ci-dessous :

- Définition de la stratégie du SI en collaboration avec la directrice du CH et en lien avec les évolutions réglementaires.

- Mise en œuvre de la stratégie : définition technique et organisationnelle, découpage en projets, rédaction de cahiers des charges et suivi de marchés publics (hors partie purement administrative), suivi de la mise en place, de la maintenance, etc.

- Gestion de l'infrastructure et des serveurs en place, et maintien en condition opérationnelle.

- Gestion des applications métier en collaboration avec les experts métier issus des différentes directions fonctionnelles et interface avec les éditeurs (sauf exceptions où le premier niveau peut être délégué à un référent métier).

- Mise à niveau et maintenance du parc de PC, clients légers, imprimantes, etc.

- Support aux utilisateurs.

Sauf quelques scripts d'ordre technique, aucun développement n'est réalisé en interne et, plus généralement, aucun développement spécifique n'est en production.

Sur le même schéma que l'informatique que nous venons de présenter, les missions du service informatique sont pratiquement les mêmes en ce qui concerne la téléphonie :

Définition et mise en application technique de la stratégie (marchés de mises en place, de maintenance, d'opérateurs, etc.), paramétrages au quotidien des autocommutateurs (2 en haute disponibilité sur le site principal et 1 situé sur l'EHPAD « La Cubissole » pour un peu plus de 600 postes) et maintien en condition opérationnelle de premier niveau en association avec l'expertise de l'intégrateur en charge de la maintenance, gestion des terminaux mobiles (GSM, DECT), etc. Le service technique assure quant à lui, la maintenance du parc de téléphones fixes.

Depuis de nombreuses années et pour la réalisation de toutes ces missions, les directions successives ont accordé leur confiance au service informatique et lui ont conféré une large autonomie.

Par ailleurs, une autre composante qui, selon les organisations et même si ce n'est pas le cas au CH du Blanc, peut se voir rattachée à la direction du système d'information : le Département d'Information Médicale (DIM) dont les missions sont « la gestion et le conseil pour le Dossier Médical du patient » [9], le codage de l'activité, « la représentation des services dans la mise en place de l'Informatique Hospitalière » [9]. Ce service est composé d'un médecin à temps partiel et de techniciens d'information médicale.

1.2.2 L'INFRASTRUCTURE TECHNIQUE ET EVOLUTIONS RECENTES

Comme nous l'évoquions en introduction, ce projet de mise en place d'une solution de haute disponibilité des serveurs fait partie d'un ensemble de projets visant à mettre à niveau et sécuriser le système informatique en vue de la mise en place du dossier médical et paramédical du patient. J'ai eu pour mission de définir cette nouvelle architecture dans son ensemble et d'en organiser les différents sous-projets. Ainsi, la majeure partie des briques techniques composant l'architecture a été revue dans le cadre de cette approche globale et est brièvement décrite à la suite de ce chapitre.

Afin d'imaginer l'ordre de grandeur de l'installation informatique, il est utile de préciser que le CH dispose, avant informatisation du dossier patient, d'environ 145 PC de bureau (très peu sont des portables).

1.2.2.1 Interconnexion des sites

Il faut distinguer ici les trois sites du Centre Hospitalier. En premier lieu, le site de la Résidence de l'Anglin est une petite structure où l'informatisation y est extrêmement faible de par son activité mais aussi par le fait que les ressources support et d'encadrement sont partagées avec le reste du CH. On n'y comptabilise donc qu'un PC relié au CH via un simple client VPN reposant sur un lien ADSL.

Le site de l'EHPAD La Cubissole est relié au CH par l'intermédiaire de pont radio 5,4 GHz répondant à la norme 802.11a et reposant sur 4 bornes MP11 de marque Proxim positionnées sur chacun des deux sites ainsi que sur un château d'eau intermédiaire (les

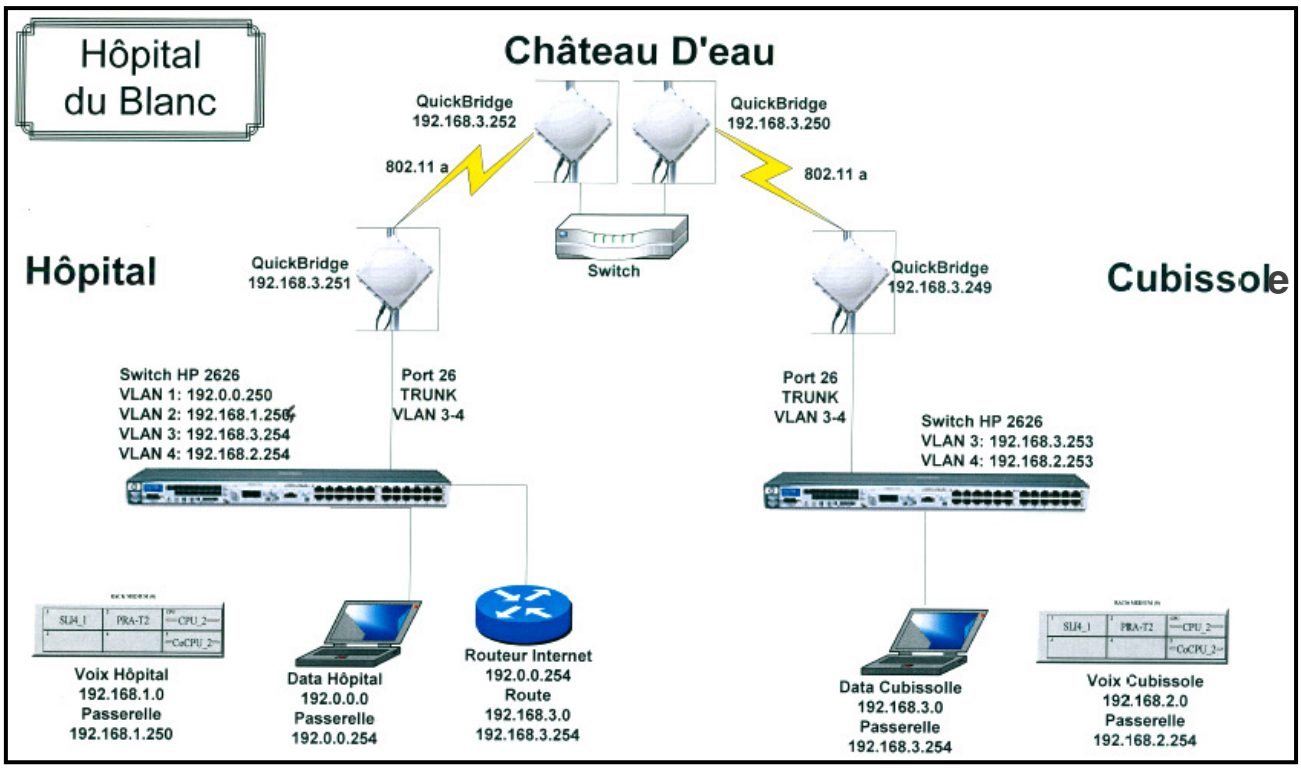


Illustration 10. Présentation schématique du lien proposé par l'intégrateur [10]

deux sites n'étant pas « à vue »). Ce lien sécurisé (Cryptage AES 128 bits) assure un débit mesuré d'environ 27 Mb/s où sont véhiculés deux réseaux distincts : l'un, prioritaire lors de l'émission, est dédié à la téléphonie IP (lien entre autocommutateurs), l'autre, permet l'échange avec les réseaux dits « utilisateurs » du CH (par opposition aux réseaux dits médicaux ou médico-techniques qui sont inexistants sur les sites distants). La partie « voie » est secourue par des lignes téléphoniques numériques de type T0 et T2 associées à des dispositifs de routage des communications.

Les accès extérieurs de l'ensemble des trois sites sont centralisés sur l'accès Internet principal du CH reposant sur la technologie SDSL dont le débit est actuellement de 2 Mb/s mais qui devrait être porté au-delà avec les projets de transmissions d'images médicales. L'ensemble des échanges vers et depuis l'extérieur ainsi que les échanges avec les réseaux médicaux et médico-techniques, se fait par l'intermédiaire d'un firewall, objet d'un marché public courant 2009.

1.2.2.2 Infrastructure réseau du CH et des sites distants

L'architecture globale est présentée à l'illustration page suivante. Le CH compte trois armoires informatiques importantes (en nombre et type de machines connectées) assurant la connexion au réseau informatique des services de courts et moyens séjours, médico-techniques (imagerie, bloc opératoire, etc.), et de consultation, etc. Ces armoires sont reliées entre elles par une boucle réseau à double attachement en fibre optique, entre les matériels actifs qui composent chaque pile. Cela permet ainsi de pallier les principales pannes potentielles (défaillance d'un connecteur optique, d'un câble optique voire, pour partie, d'un élément actif de la pile). Le cœur de réseau n'est, à ce jour pas doublé, mais cela pourra faire l'objet d'une étude ultérieure.

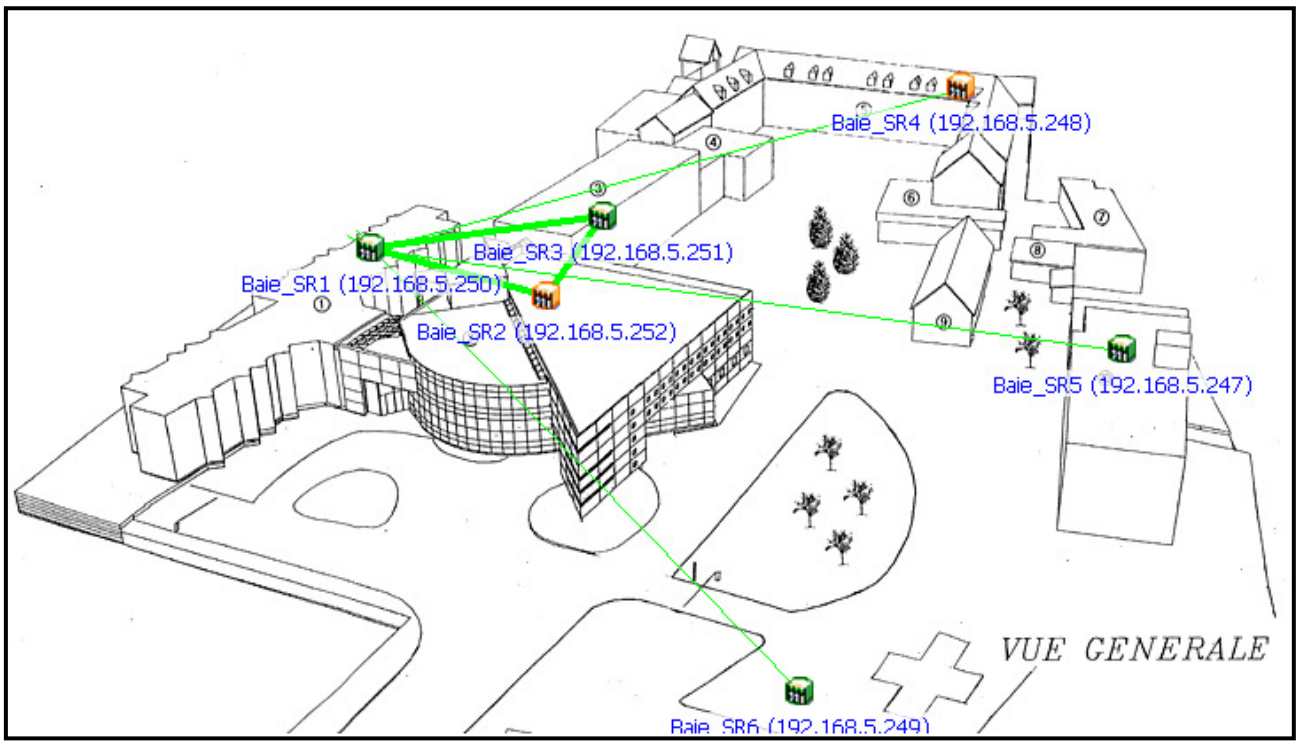


Illustration 11. Vue synoptique de l'implantation géographique des armoires informatiques du CH du Blanc et leurs interconnexions [11]

Les trois autres armoires sont, quant à elles, reliées par des liens fibre optique simples.

Ces liens reposent sur des fibres optiques dont les époques et les spécifications divergent : câbles de 6 à 12 brins optiques de type 50/125 ou 62.5/125 essentiellement multimodes.

Le câblage Ethernet, précédemment réalisé sans réel respect des normes, a été contrôlé et mis à niveau quand cela a été nécessaire pour qu'il ne reste, sur le site du CH, qu'au minimum des liens de catégorie 5e. Malheureusement, cette opération n'a pas pu encore être réalisée sur le site de la Cubissole où manifestement se posent, au moins, des problèmes de longueurs de câble et de connectique d'extrémité. Pour cette réfection du câblage et afin de répondre aux nouvelles demandes, un marché public dit « à bons de commande » que j'ai eu pour mission d'écrire et de suivre, a été mis en place. Ces opérations s'apparentant essentiellement à des opérations de travaux (contraintes par les

règles d'hygiène hospitalière), leur transfert au service technique était alors un objectif, depuis atteint.

L'ensemble du matériel actif a, quant à lui, été revu pour l'ensemble du site du CH en début 2010, concrétisant les aspects que nous venons d'aborder (boucle réseau, liens croisés), ainsi que la segmentation des réseaux et la bascule progressive du réseau « utilisateurs » vers un adressage privé [12]. S'agissant d'un élément stratégique pour le CH du Blanc, j'ai intégré à la rédaction du marché, sa couverture par un contrat de maintenance.

Finalement, un réseau Wifi couvre les circulations et les salles de soins des services de court et moyen séjour ainsi que l'ensemble du plateau technique en prévision du déploiement du dossier patient. Il ne sera utilisé que pour l'accès au serveur (via une connexion de type « bureau à distance » reposant sur le protocole RDP). Aucune solution dédiée aux visiteurs (patients, personnels remplaçants, etc.) ne leur a, pour le moment, été offerte.

1.2.2.3 Les applications informatiques

Après consolidation en 2008, l'hôpital du Blanc est équipé d'un Progiciel de Gestion Intégrée (PGI) hospitalier de l'éditeur Berger-Levrault couvrant les domaines de la gestion administrative des patients et des séjours, de la facturation, des dépenses, de la paie, de la comptabilité, de la gestion analytique, des stocks, etc. Précédemment équipé de deux solutions différentes pour, d'une part, la partie recette et gestion administrative des patients et pour, d'autre part, tous les autres domaines évoqués ci-dessus, le CH a dû faire face à des problèmes de qualité, de surcoût et de support de la première solution et, plus globalement, à une absence totale d'intégration. Cette situation a conduit la direction du CH à me confier l'assainissement des applications informatiques qui a, en premier lieu,

entrainé ce nouveau changement au 1er Janvier 2008 retardant d'autant la décision de planifier l'informatisation du dossier médical et paramédical du patient et de tous les projets connexes ici, en partie, évoqués. Cette suite logicielle ainsi assainie pouvait alors pleinement jouer son rôle de socle pour le dossier patient car, au-delà des structures, s'y trouvent gérés toutes les identités, les mouvements puis, en sortie, tous les aspects liés à la facturation du séjour et à la comptabilisation de l'activité. Un socle défaillant aurait présenté un risque important tant de par sa probabilité que de par ses conséquences sur l'utilisation du DPI.

Ce changement a été l'occasion de faire se reposer les nombreux échanges internes et externes sur une solution globale de type : EAI (envoi des identités et retours des actes

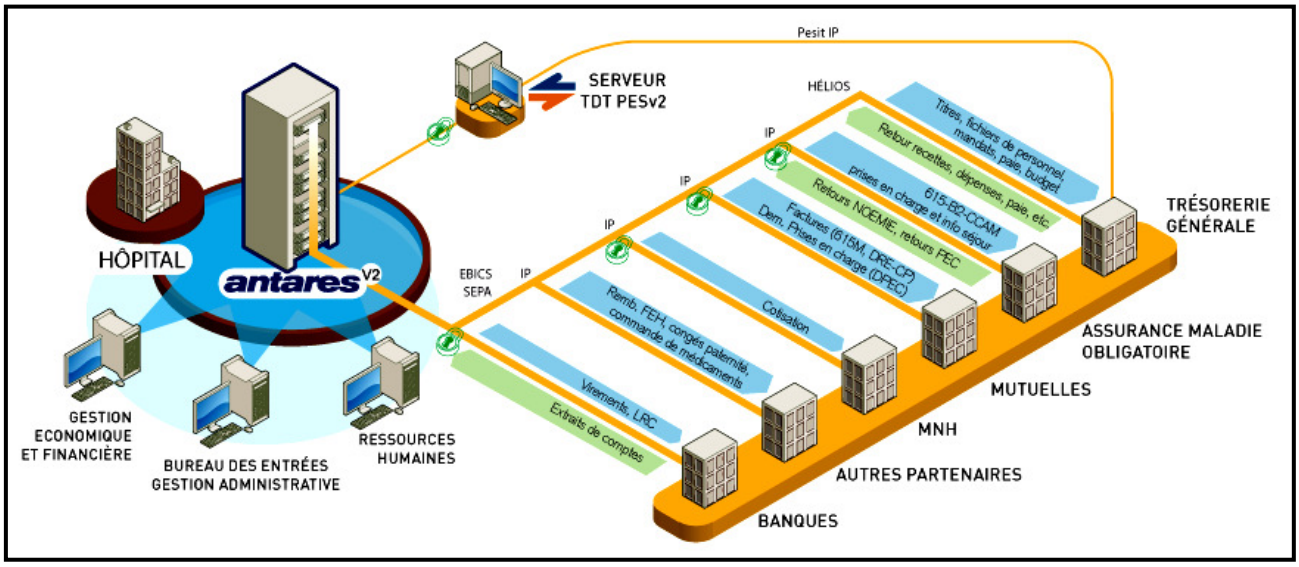


Illustration 12. Vue des principaux flux administratifs et financiers offerts par la solution d'EDI du CH [13]

vers le progiciel du service d'imagerie) et EDI (voir détails des correspondants et des contenus à l'Illustration 12). L'objectif étant de venir y adjoindre les futurs flux internes ou externes liés au dossier patient :

1. Identités et retour de résultats de laboratoire,

2. Eventuellement identités vers des matériels médicaux,

3. Création et alimentation du DMP (dossier médical de synthèse national),

4. Identités et mouvements vers des applications diverses telles que celles liées au système de taxation téléphonique hospitalière ou encore au système de gestion de commandes de repas pour les patients et accompagnants, etc.).

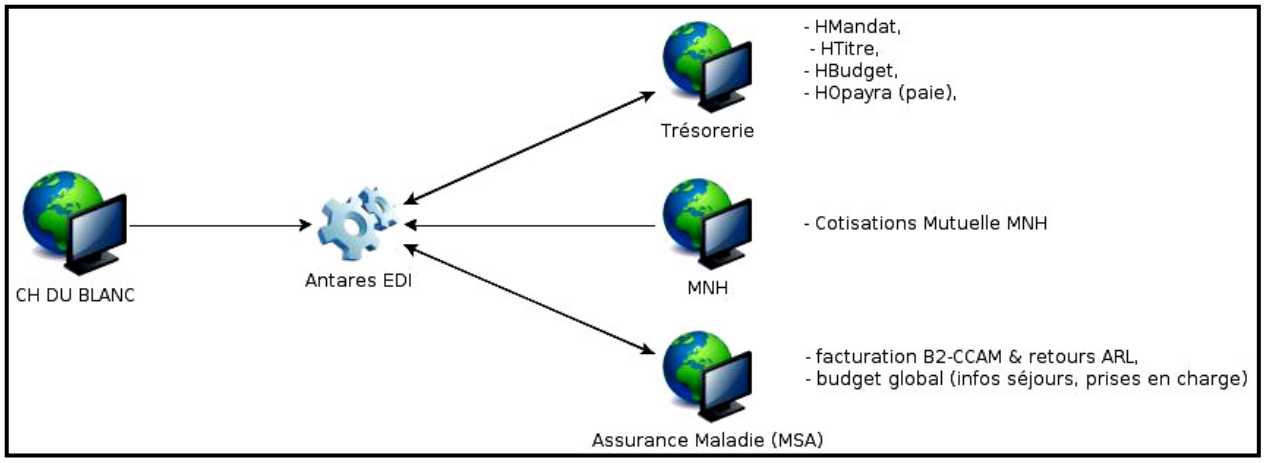


Illustration 13. Echanges informatisés mis en place lors des premières phases

Au début de ce projet, donc avant la mise en place du DPI, les applications métier notables utilisées sur le CH, hors applications médicales, techniques (code barres, consoles d'administrations, etc.) ou de bureautique, sont les suivantes :

Tableau I : Applications métier principales avant projet

Editeur / Application ou Gamme	Descriptif
Berger-Levrault / DIS (gamme)	PGI (GAP, recettes, Dépenses, Stocks, etc.)
Enovacom / Antares v2	Plateforme EAI / EDI
Medasys / Sirilog 6.8	Système d'Information de Radiologie (RIS)
Yes / Gala (gamme)	Système documentaire et gestion qualité
Aidel / Superdoc	Système de gestion documentaire des ressources du CDI de l'IFSI.
FSI / Winrest	Gestion des commandes de repas
PMSIPilot / Gamme PmsiPilot	Analyse activité hospitalière, contrôle de gestion, qualité facturation/codage, etc.
Axilog / InfanSoft V2	Dossier patient spécifique à la pédiatrie

Editeur / Application ou Gamme	Descriptif
Vidal / Vidal Expert	Données et Applications pour la consultation ou l'intégration à d'autres outils (via WebServices) de la base de données médicamenteuse.
MGDIS / SOFI Analyse Financière	Applications d'analyse financière (planification pluriannuelle des investissements et de l'exploitation ainsi que gestion financière rétrospective et prospective).
ATIH / Fichsup, Genrha, Magic, Preface, Genrsa, etc.	Diverses applications liées à la déclaration de l'activité.
Interbat / Marchés Sécurisés	Gestion électronique des marchés publics (de la publication à l'ouverture)
Cefotec / Genepi	Gestion de la scolarité (IFSI/IFAS)
Etc.	

Leur nombre restreint vient de la recherche de solutions au maximum intégrées qui apparaissent souvent plus compatibles avec la taille de la structure (peu de moyens pour les financer, peu de ressources informatiques pour les gérer, des profils polyvalents plus que des spécialistes métier pour les exploiter, etc.). Ne sont pas non plus listées ici les quelques applications conservées pour archive.

1.2.2.4 Les serveurs

Différents serveurs permettent le fonctionnement des applications. Le choix du CH aura souvent été la multiplication de petits serveurs afin d'isoler les applications les unes des autres. Il en aura résulté une plus grande souplesse d'adaptation des environnements aux progiciels qu'ils hébergeaient mais, en contrepartie, un plus grand nombre de machines et de sauvegardes à gérer. Toujours hors médical, on comptabilise les serveurs suivants :

Tableau II : Serveurs en production avant projet

Serveurs et caractéristiques	Rôles
2 serveurs HP en cluster (UNIX) répartis entre SR1 et SR2 avec sauvegarde sur DAT.	Partie recette et gestion patient AGFA conservée pour archive.
Serveur IBM (Unix SCO) avec sauvegarde sur DAT	Complément de la partie AGFA pour ce que couvrira par la suite le PGI géré avec les produits Berger-Levrault (comptabilisé, paie, finance, stocks, etc.).
Serveur HP ML350 (x86) équipé de Windows 2000 avec sauvegarde sur DAT.	Système d'Information de Radiologie (RIS)
Serveur HP ML110 (x86) équipé de Windows 2003 avec sauvegarde sur DAT mais dépourvu de dispositif de sécurité tel que RAID, double alimentation, etc.	Système documentaire et gestion qualité
Serveur DELL PowerEdge 1950 III équipé de Windows 2003 sans dispositif de sauvegarde intégré (export).	Gestion des commandes de repas Winrest
Serveur DELL PowerEdge 850 équipé de Linux Mandriva 2009.0 non sauvegardé (son rôle se limite au retraitement des données normalisées transmises aux tutelles donc régénérables à volonté).	Analyse activité hospitalière, contrôle de gestion, qualité facturation/codage, etc.
Serveur Digital (Unix) avec sauvegarde sur DAT.	Ancien PGI SYMPHONIE conservé uniquement pour historique.
Serveur IBM x3250 équipé de Windows 2003 sans dispositif de sauvegarde intégré (export)	Serveur EAI / EDI (Enovacom / Antares v2)
Serveur HP ML110 (x86) équipé de Linux RedHat Enterprise Release 4 avec sauvegarde sur LTO1 mais dépourvu double alimentation, etc.	Contrôleur de domaine
Serveur Maxdata Platinium 300 IR M8 équipé de Windows 2003 sans dispositif de sauvegarde intégré (export)	Gestion centralisée de l'antivirus.
PC de bureau « simple »	Système de taxation hospitalière
PC de bureau équipé de RAID, etc.	Système d'information dédié pédiatrie

Un contrat de maintenance global, que j'ai pu mettre en place au travers d'un marché public, aura permis à partir de 2008 de protéger ces matériels souvent vieillissant (hors cluster HP dont le besoin ne justifiait pas le coût). Quant à la multitude des supports de sauvegarde, ils sont depuis quelques années stockés dans un coffre ignifugé positionné dans un autre bâtiment et appartenant à une « zone de feu » différente.

L'absence d'onduleur restait un point important à traiter de par le risque qu'elle fait encourir au système. Il s'agit d'un point de défaillance unique [14] que nous nous devions d'éliminer.

1.2.2.5 Divers dispositifs

Différents dispositifs mis en place au cours des dernières années complètent l'installation parmi lesquels nous pouvons citer un firewall qui assure l'interconnexion sécurisée entre deux des sites du CH, avec d'autres CH ou encore entre le réseau « utilisateurs » et les réseaux médicaux. Il permet aussi les télémaintenances sur certains matériels, etc.

Une gestion centralisée de l'antivirus ou encore des procédures permettant une plus grande fiabilité ainsi qu'une meilleure maintenabilité ont été mises en œuvre.

1.3 PROBLEMATIQUE

L'objectif du centre hospitalier confié à la direction du système d'information est la mise en place d'un système d'information de santé permettant aux personnels médicaux et paramédicaux de mener à bien la prise en charge du patient. La problématique de mise en place de « serveurs » (terme très générique pour la direction du CH comme pour les utilisateurs) dits « sécurisés » (terme tout aussi générique) sera donc l'objet du projet que nous développerons dans ce document.

1.3.1 PRESENTATION DU CONTEXTE DU PROJET

Le dossier patient informatisé allait donc permettre d'assurer une meilleure traçabilité des interventions de chacun des professionnels notamment dans le cadre du « contrat de bon usage du médicament » [1] dans lequel l'établissement s'est engagé. Mais, il allait surtout permettre aux différents professionnels de santé d'en exploiter des informations (résultats d'analyses biologiques automatiquement intégrés, constantes saisies par les infirmières, prescriptions en tous genres saisies par les médecins, courbes d'évolution, etc.), d'échanger dans le but d'assurer la continuité de la prise en charge du patient (au travers des transmissions ciblées infirmières notamment, etc.). Et cela, d'une manière qui se voulait plus efficace, faisant ressortir les anomalies (de prescriptions contre-indiquées, de données vitales hors normes, d'actions prévues mais non encore réalisées, de traitements arrivant à échéance, etc.). L'outil informatique se doit donc, au-delà d'un simple changement de support, de proposer une réelle valeur ajoutée aux utilisateurs. Celle-ci devant surpasser la surcharge induite par l'appropriation d'un nouvel outil ou tout simplement l'appropriation de l'informatique qui est pour beaucoup un cap important.

Cet outil, déployé dans les services selon des modalités et un planning qui seront définis par le comité de pilotage DPI dont j'aurais en charge l'animation, sera diffusé essentiellement au travers de clients légers fixes et mobiles (« qui coûtent en général moins cher, durent plus longtemps et consomment moins d'énergie » [15]). Pour des raisons de fiabilité, de maintenabilité, etc., le choix a été fait de rendre les clients légers les plus simples et les plus standards possibles et de déporter sur un serveur dédié, l'ensemble de la partie traitement de l'application choisie (exécutant par exemple Microsoft Terminal Server, Citrix XenApp, Systancia AppliDis ou autre solution du marché) ainsi que toute la gestion des périphériques d'impression. Cette problématique devra être traitée

dès le choix de la solution logicielle opéré et ce dans le cadre la mise en place la plateforme nécessaire (matériels et logiciels).

Avant d'aborder le choix quant au niveau de disponibilité recherché, il apparaît important de préciser ici, que le niveau d'exigence des utilisateurs apparaît comme élevé. Ceux-ci ont soulevé le problème dès la phase d'étude des solutions logicielles disponibles sur le marché. La simple évocation de coupures régulières pour les mises à jour applicatives est à elle seule source d'inquiétude.

Finalement, compte-tenu du délai réduit entre la décision d'informatiser le dossier patient et les engagements du CH dans le cadre du contrat de bon usage du médicament, du mode de choix de la solution logicielle (marché public de type « dialogue compétitif » donc procédure longue), de la compatibilité nécessaire de la solution matérielle avec la solution applicative (et non l'inverse), plannings et ordonnancement des tâches auront été stratégiques. Nous pourrons constater un choix rapide suivi de la mise en place tout aussi rapide des éléments nécessaires au DPI puis un rythme beaucoup plus lent concernant les éléments annexes qui ne revêtaient pas un caractère d'urgence.

1.3.2 DEFINITION DU PROJET

1.3.2.1 Expressions des besoins des utilisateurs et décideurs

Après échanges avec les différents acteurs (utilisateurs et direction qui sont ce qu'Alain Le Put ou encore Patrice Besse [16] qualifient de clients dont aucun des points de vue ne doit être négligé), il en est ressorti les objectifs suivants (vision de l'utilisateur) :

- Continuité de service : il s'est révélé que le temps de coupure maximum toléré n'était pas une valeur clairement définie mais une fourchette dont les bornes pouvaient bouger en fonction de la contrainte de coût qui allait apparaître tant pour

la partie investissement qu'exploitation. L'objectif décidé fut qu'une coupure générale imprévue ne dure pas plus d'une heure, idéalement 15 minutes, mais que cette valeur pourrait être portée à trois heures en fonctions des éléments que nous venons de citer. Il est à noter qu'aucun coût lié à un arrêt de l'activité n'a été évalué car, au-delà du fait qu'il ne s'agisse pas ici de production industrielle, l'impact est surtout lié à la sécurité lors de la prise en charge du patient.

- Le système devait pouvoir être autonome dans la gestion des pannes liées au matériel c'est-à-dire capable de redémarrer les services nécessaires au bon fonctionnement du dossier patient sur une éventuelle autre machine ou toute solution reposant sur ce principe.

- Le système devait être pleinement compatible avec la solution choisie, performant et être facilement évolutif pour s'adapter au volume de données croissant et à l'inévitable évolution logicielle.

- Perte de données maximum : en cas de problème majeur, la perte de données pour un retour à un état précédent stable ne devrait pas excéder 24h mais était attendue une solution ne permettant de ne perdre au maximum que la saisie en cours.

- Cette solution visant à protéger divers serveurs devra rester économiquement supportable par le CH.

Au-delà et de manière moins stratégique, il serait apprécié que cette solution permette de sécuriser les autres applications du CH avec, dans la mesure du possible, une possibilité de redémarrage automatique.

Le service informatique envisage de profiter de ce changement pour changer le contrôleur de domaine. Il s'agit d'une machine supportant un système Linux RedHat complété de

Samba. Cette machine est ancienne (2003), peu sécurisée alors que son rôle est devenu stratégique, et l'espace disque vient à manquer.

Au-delà des échanges électroniques et autres comptes-rendus de réunion, le principal écrit, nécessaire à la contractualisation des objectifs avec les décideurs [14], aura été le cahier des charges soumis aux entreprises puis le choix d'un prestataire et par conséquent de la solution qu'il propose. Un formaliste plus rigoureux et surtout associant les principaux utilisateurs (référents) aurait probablement eu une légitimité supérieure dans l'hypothèse d'une panne longue n'excédant pas forcément les objectifs initiaux. Cela constitue, à mon sens, un point d'amélioration.

Finalement, par le simple fait des coupures organisées dans le cadre des mises à jours des systèmes mais surtout des applicatifs, il apparaît évident que le taux de disponibilité annuelle "réel" pourrait se situer seulement autour de 99,9% soit presque 9h de coupure par an. Néanmoins, comme nous l'indique René J Chevance, « si le système, dans ses spécifications, présente des périodes d'arrêt programmé, ces périodes d'arrêt ne sont pas à prendre en compte dans les calculs : seules les périodes d'arrêt non programmé sont à considérer » [14]. « Ces phases d'arrêt programmé ne sont pas considérées comme des périodes d'indisponibilité » [14]. Tel est bien notre cas ici.

1.3.2.2 Etude d'impact

Réduit à l'aspect serveurs et stockage associé (ce qui correspond au périmètre du projet ici présenté), il était nécessaire de mesurer l'impact de cette décision sur l'environnement actuel et futur (pour le DPI) afin de définir le périmètre que devait couvrir cette nouvelle étape de sécurisation. L'étude des processus aura été indispensable.

Le tableau suivant reprend les machines impactées (machine existantes et machines attendues dans le cadre de cette mise en place), leurs rôles vis-à-vis de l'exploitation du dossier patient, et un critère d'importance associé. Nous nous intéressons ici au choix des composants à rendre disponibles et pour lesquels « il convient d'adopter un traitement différencié pour chacun de ses constituants en fonction de son degré de criticité dans le système » [14]. Comme nous le rappelle Taneja Group, « il suffit qu'un serveur ou un fichier essentiel soit indisponible, même quelques heures, pour que les conséquences soient considérables » [17].

Tableau III : Machines liées au DPI et impactées par le projet, rôles et importances

Machine	Environnement	Rôles	Importance
DIS	Serveur Linux (RedHat Enterprise Linux Server release 5.3) proposé par l'éditeur retenu pour le DPI	Données et quelques traitements pour le DPI.	Indispensable
SRVRDS1	Serveur Windows 2008 R2 Terminal Server	Diffusion de l'application aux clients légers des services de soin et impressions.	Indispensable
ZEUS	Serveur Windows 2008	Gestions des stratégies, des scripts d'ouverture de session, des fichiers utilisateurs, etc.	Indispensable
SRVENOVEAI	Serveur Windows 2003	Plateforme EAI / EDI importante pour les échanges liés aux résultats d'analyses biologiques voire avec le RIS (imagerie).	Haute.
Rappel des besoins fonctionnels : coupure maximum 1h à 3h avec objectif 15 minutes, redémarrage automatique en cas de panne. En cas de crash majeur : perte de données maximum de 24h.			

Il est à noter qu'il n'y a pas de serveur d'affectation dynamique d'adresses (DHCP) ou autre serveur nécessaire au fonctionnement du DPI au CH du Blanc. Il est aussi à noter

que le serveur Linux RedHat, correspond à l'environnement proposé par l'éditeur pour y installer les applications nécessaires au dossier patient mais aussi pour migrer, depuis un serveur UNIX SCO 5.0.7, celles liées à l'ensemble du Progiciel de Gestion Intégrée assurant, entre autres, la gestion administrative du patient (indispensable pour la gestion des identités et mouvements), la paie, la gestion économique, etc.

Dans la mesure du possible, il était souhaitable et dans certains cas indispensable de sécuriser voire augmenter les performances des machines suivantes d'autant que certaines ne sont pas très récentes voire anciennes (cas de la machine Sirilog de 2003 ou encore SrvQHS de 2005) ou ne sont pas sécurisées (cas de la machine SRVQHS et PMSIPilot qui ne disposent pas de RAID, ni double alimentation, etc.). :

Tableau IV : Machines supplémentaires auxquelles le projet devrait bénéficier

Machine	Environnement	Rôles
SIRILOG	Windows 2000	Système d'Information de Radiologie (RIS)
SRVQHS	Windows 2003	Système documentaire et gestion qualité (disparition à terme des « procédures papier »)
SRVWINREST	Windows 2003	Gestion des commandes de repas patients Winrest
PMSIPILOT	LINUX Mandriva 2009.0	Analyse activité hospitalière, contrôle de gestion, qualité facturation/codage, etc.
SRVVIRUS	Windows 2003	Gestion centralisée de l'antivirus
Rappel des besoins fonctionnels : Sécuriser ces machines et accroitre leur performance respective. Une coupure plus longue que pour les machines liées au DPI est tolérée, un redémarrage automatique serait apprécié mais pas indispensable. En cas de crash majeur : perte de données maximum de 24h pour Sirilog et sans objectifs précis pour les autres machines.		

1.3.2.3 Choix technologiques

Au vu des éléments ayant permis de déterminer les risques, les activités critiques et les systèmes informatiques indispensables à ces activités, il est apparu nécessaire de mettre

en place une solution en haute disponibilité pour au moins quatre serveurs (mais idéalement plus) afin de pouvoir :

- pallier, de manière automatique, la panne d'un élément matériel où d'un serveur complet, voire à pallier la disparition (même définitive), au moins en mode dégradé et sans intervention humaine, d'une des deux salles informatiques.

- pallier, dans la mesure du possible, un plantage système.

- permettre de retrouver un système stable après un plantage applicatif ou système majeur en quelques dizaines de minutes.

- limiter au maximum les coupures liées à une maintenance matérielle programmée.

L'architecture se devait aussi d'être évolutive afin de pouvoir, par exemple, doubler les capacités d'utilisation (en nombre d'utilisateurs connectés) des applications liées au DPI.

Problème : la consolidation de cet ensemble de serveurs sur une machine dite sécurisée n'est pas envisageable en l'état, du fait des différents systèmes d'exploitation (Linux et Windows), et des restrictions quant à la possibilité d'installation de toutes les applications sur une même version des systèmes d'exploitation (Windows 2000, 2003, 2008, 2008 R2) ou tout simplement dans un environnement logiciel commun et partagé. Quant à l'idée de doubler chacune des machines importantes, même en se limitant aux machines stratégiques, elle n'était économiquement pas envisageable : « Les solutions de continuité d'activité traditionnelles sont onéreuses et complexes à déployer, ce qui les rend inabordables pour la plupart des PME » [18].

La seconde approche visant à externaliser les serveurs n'a, dès le départ, pas été retenue. Elle aurait été plus audacieuse puisque moins admise voire conventionnelle

qu'aujourd'hui, mais les principaux freins dans le cas du CH du Blanc auraient sans doute été les coupures fréquentes des liens opérateurs et la nécessité de préparer l'architecture logicielle aux contraintes de débit, etc.

Finalement, largement promue par les médias spécialisés ou les équipes commerciales des SSII, « la virtualisation est devenue un outil majeur de modernisation […] car elle permet d'optimiser l'utilisation des ressources […] et d'améliorer la résilience » [19]. Ce concept présenté comme une solution incontournable à la problématique du CH du Blanc devait être étudié. D'autant que, selon SearchDataCenter.com, parmi les utilisations les plus populaires, on retrouve « l'amélioration de la disponibilité des applications et des capacités de reprises après désastre » [19]. « En effet, un choix initial d'investissement dans une infrastructure de serveurs, stockage et communications réseau redondante pourra bénéficier en termes de disponibilité à l'ensemble des serveurs virtuels qui seront mis en œuvre » [20].

Il devenait ainsi possible de consolider, sur cette infrastructure matérielle et logicielle la plupart des serveurs du CH. « Les machines virtuelles [...] constituent des charges de travail encapsulées et protégées (faisant office de serveurs indépendants du matériel) [Permettant ainsi de] consolider en toute sécurité plus d'applications sur moins de serveurs, sans vous soucier d'éventuels conflits. » [17] répondant ainsi à la contrainte liée à la multitude des environnements d'exécution. De plus, l'ajout ou le remplacement de pratiquement l'ensemble des serveurs (vieillissants) via une solution reposant sur la virtualisation « permet de réduire les coûts par une réduction du nombre de machines physiques et donc une meilleure utilisation des matériels, une moindre consommation électrique, des gains de place dans les salles machines, etc. » [21].

« Initialement appréciée en tant que technologie de consolidation des serveurs et de réduction des coûts, la virtualisation devient progressivement la plateforme privilégiée par l'industrie pour le déploiement de solutions BC/DR » [17]. Encore, « selon les travaux de Taneja Group, la volonté d'améliorer leurs opérations BC/DR constitue l'une des deux principales raisons pour lesquelles les PME adoptent la virtualisation » [17]. En cas de crash système ou applicatif majeur, il devient alors possible de repartir d'un environnement stable précédemment sauvegardé et testé. Tout comme ce système surveille en permanence l'état des machines et de leur système d'exploitation. En cas de défaillance, il déplace et redémarre automatiquement les machines virtuelles sur un autre serveur physique [17], répondant ainsi à une problématique précédemment soulevée.

Quant au stockage, il « représente le deuxième pilier de l'informatique, car c'est là que résident les données. Les fournisseurs de stockage ont développé des offres de plus en plus indépendantes des serveurs proposant des fonctions très intéressantes » [22]. Plus globalement, de nombreux articles ou autres documents, même partiaux, dits de conseil aux PME, mettent en avant « une architecture de stockage de type SAN (même d'entrée de gamme) indispensable pour bénéficier des aspects de haute disponibilité ou de mise en œuvre de PRA » [21]

Mais cette technologie peut-elle être présentée comme un standard dont la pérennité ne semble pas compromise ? En réalité, la virtualisation n'est pas une technologie nouvelle et on note une expansion de son utilisation. En effet, selon SearchDataCenter.com, « 60% des entreprises l'ont largement déployée » [19] dont « 43% dans le cadre d'un PRA ou d'une solution de haute disponibilité » [19]. « La moitié des charges de travail des serveurs du monde entier sera virtualisée d'ici 2012, selon le cabinet d'analyse Gartner Inc. De

plus, Gartner prévoit que d'ici 2016, 80 % des charges de travail prises en charge par un équipement x86 s'exécuteront sur des machines virtuelles » [15].

Voyons maintenant ce qu'il est possible d'attendre de ces technologies.

2. ETAT DE L'ART

La solution mise en place repose sur deux concepts indépendants mais complémentaires voire indissociables, dans le projet ici évoqué, que sont : la haute disponibilité et la virtualisation. Tout d'abord, « un système à haute disponibilité doit masquer les défaillances auxquelles il est sujet, vis-à-vis de son environnement » [14]. Il apporte des réponses aux conséquences de pannes d'un élément constituant l'architecture ou encore, dans le cas du CH du Blanc, à l'indisponibilité totale de l'une des deux salles informatiques. C'est le socle technique nécessaire de la disponibilité attendue des applications (limité au périmètre qui lui incombe). La virtualisation se présente, quant à elle, comme un élément majeur de la consolidation des serveurs tout en en assurant l'indépendance des systèmes hébergés avec le matériel mais aussi entre chacun des systèmes virtualisés. L'association de ces deux technologies que nous allons voir permet de mettre à disposition de multiples environnements virtualisés, les capacités de continuité de service de l'infrastructure.

2.1 LA HAUTE DISPONIBILIITE

« La continuité d'activité est une affaire d'organisation, de planification et de technologie » [22]. En effet « toute organisation doit donc mener des actions visant à prendre conscience de son environnement et à comprendre son propre fonctionnement. Ce n'est qu'à cette condition qu'elle aura en main les paramètres lui permettant de maîtriser sa continuité » [22]. Tous les aspects qui précèdent ou entourent la mise en place de l'infrastructure technique comme l'identification et la caractérisation des menaces mais aussi le choix quant à celles retenues dans le périmètre que devra couvrir la solution

choisie font l'objet de nombreux ouvrages et ne sont pas directement l'objet de ce chapitre qui sera consacré aux concepts techniques permettant d'atteindre les objectifs.

Avant de définir ce concept de haute disponibilité, il apparaît intéressant de rappeler que selon le Gartner Group [23], les causes d'indisponibilité imprévue d'une ressource logicielle quelconque (par exemple module de prescription informatisée dans notre cas), sont majoritairement la conséquence d'erreurs humaines (40%) ou d'erreurs logicielles (40%). Par conséquent, la mise en place d'une plateforme de haute disponibilité telle que nous l'abordons au cours du présent document n'aura une influence sur la continuité d'activité, en dehors des fonctionnalités de retour à un état stable antérieur, qu'au maximum sur les 20% des causes restantes et limitée au périmètre du projet (qui, par exemple, ne tient pas compte des aspects relevant des projets dits réseau et traités par ailleurs). Bien évidemment, les conséquences nous obligent à traiter sur un plan plus large l'ensemble des catégories de causes et pour chacune d'elles, identifier, évaluer et inclure ou non dans le périmètre retenu les menaces jugées pertinentes en fonction de leurs conséquences (humaines, financières, etc.) et de leur probabilité d'occurrence. Emmanuel Besluau au premier chapitre de « Management de la Continuité d'Activité » [22], nous expose de manière assez pragmatique et opérationnelle quelques méthodes et outils (cartographie des risques, arbres de défaillance, etc.) permettant de conduire cette

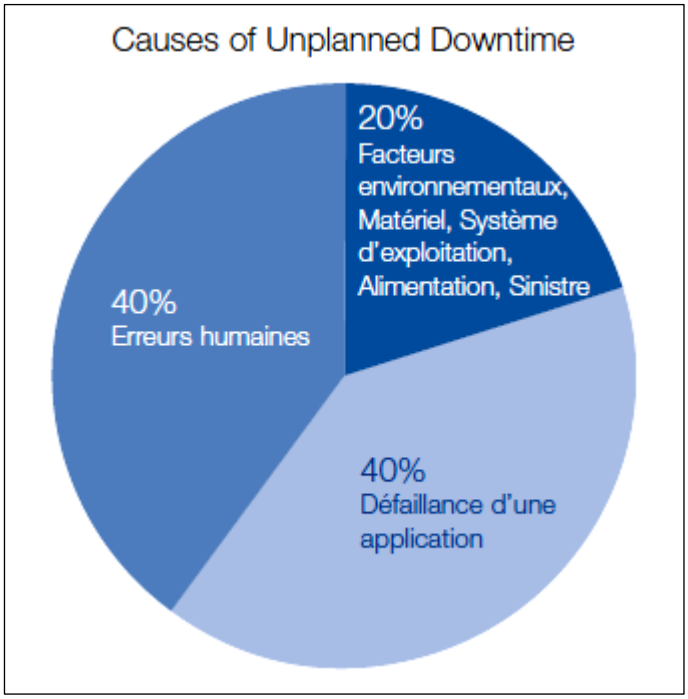


Illustration 14. Causes of unplanned Downtime [23]

démarche. Celle-ci sera alors le point de départ de l'ensemble du projet (voire des sous-projets) de continuité d'activité.

2.1.1 CONCEPT DE DISPONIBILITE ET GRANDEURS CARACTERISTIQUES

« La disponibilité d'une machine indique la proportion du temps pendant lequel cette machine fonctionne comme prévu » [22]. Il s'agit de la « capacité à assurer le service spécifié et le degré de confiance que l'utilisateur peut accorder aux services rendus par les systèmes informatiques » [14]. « Les systèmes informatiques étant sujets aux défaillances, les concepteurs de ces systèmes ont recherché les moyens de pallier ces défaillances. » [14]. Néanmoins, il apparaît important de rappeler ici que, bien que la prise en compte du système doive se faire de manière globale « il convient d'adopter un traitement différencié pour chacun de ses constituants en fonction de son degré de criticité dans le système » [14] et que, comme présenté au chapitre premier comme une contrainte forte, inhérente à tous projets mais à celui-ci en particulier, la mise en perspective d'un objectif et des solutions techniques qui permettent d'y répondre « se traduit par un compromis entre l'atteinte de l'objectif et le coût des solutions » [22].

2.1.1.1 Mesure de la disponibilité et grandeurs caractéristiques

Pour mesurer la disponibilité, on classe généralement les niveaux de disponibilité selon leur nombre de "9" à ne pas dépasser pour respecter les temps d'arrêt [22] [14] :

Tableau V : Classification des systèmes suivant leur disponibilité (et temps maximum d'arrêt) [14]

Classe de 9	Disponibilité	Indisponibilité maximum par an en minutes	Type de système
1	90 %	50 000	Non géré (unmanaged)
2	99 %	5 000	Géré (managed)
3	99,9 %	500	Bien géré (well managed)

Classe de 9	Disponibilité	Indisponibilité maximum par an en minutes	Type de système
4	99,99 %	50	Tolérant les fautes (Fault-tolerant)
5	99,999 %	5	Haute disponibilité (High Availability)
6	99,9999 %	0,5	Très haute disponibilité (Very High Availability)
7	99,99999 %	0,05	Ultra haute disponibilité (Ultra Availability)

Cette disponibilité (ou Availability [14]) est souvent mesurée en année pleine mais aussi en moyenne annuelle sur 5 ans [22]. On peut observer que ces temps apparaissent comme très réduits pour la remise en marche ou le remplacement d'un composant. « Il faut donc analyser la disponibilité sous ses deux constituants : la panne et la facilité de réparation » [22] (ou maintenabilité [14]).

2.1.1.2 Chronologie d'une défaillance et grandeurs caractéristiques

Le schéma suivant (Illustration 15) présente de manière synthétique et chronologique, les différentes phases qui constituent de fait des critères permettant de mesurer, en cas de sinistre, l'indisponibilité du système et les conséquences induites (la perte de données, etc.) ainsi que les leviers sur lesquels il est possible d'agir en amont pour réduire le délai jusqu'au retour à une situation normale.

La flèche B indique le moment du sinistre. Pour le cas où il s'avérerait nécessaire de procéder à une restauration de données, le RPO représente la durée pendant laquelle le fonctionnement du système était normal mais pour laquelle les données ne pourraient pas être récupérées. Les dispositifs limitant le recours aux sauvegardes (duplication des données, etc.) et la fréquence des sauvegardes seront donc des points stratégiques que nous aborderons ci-après. Le RTO (« Temps de récupération cible » [22]) met, quant à lui, en exergue l'importance de la maintenabilité. Nous verrons que la redondance du matériel

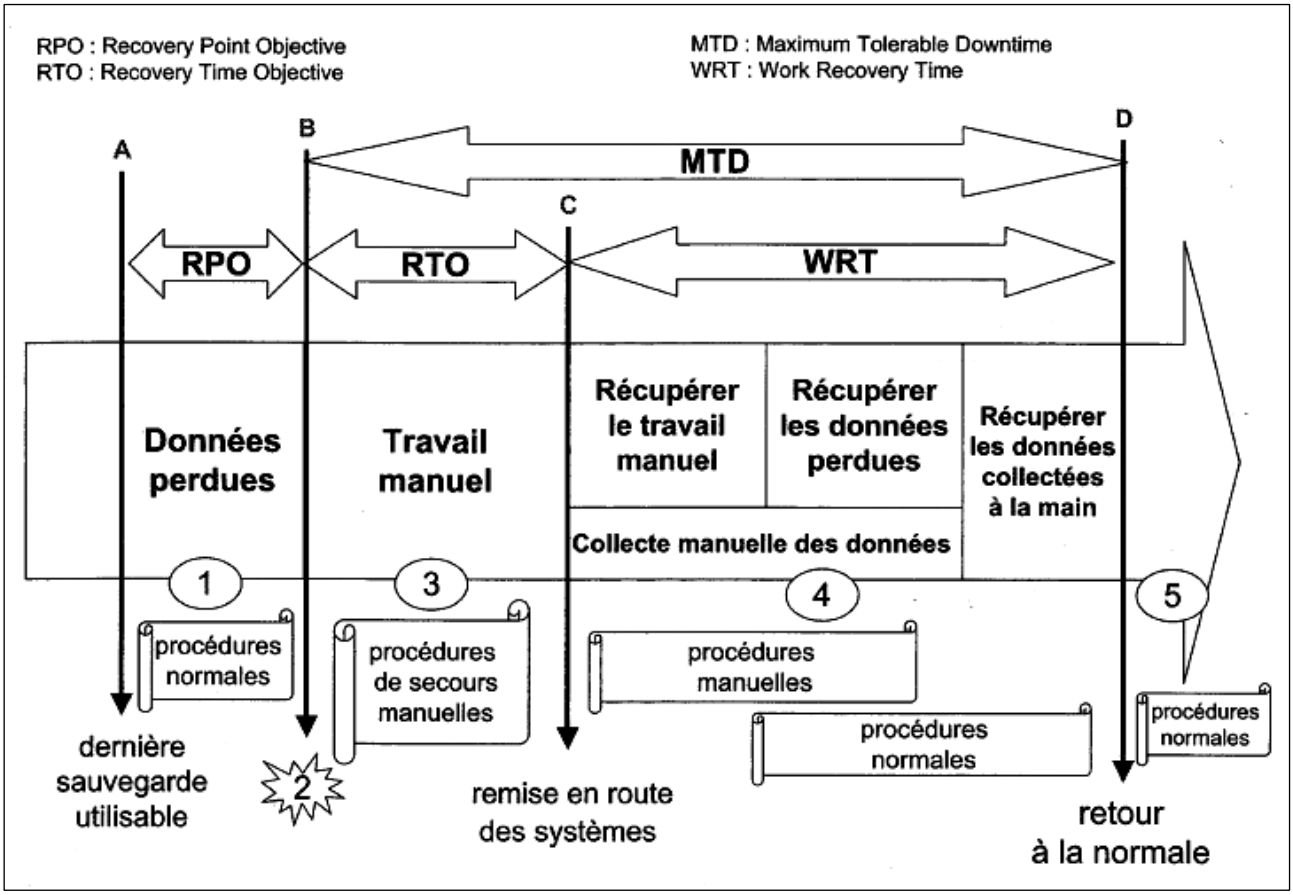


Illustration 15. Déroulement d'un sinistre et impact sur les activités [22]

couplée à la possibilité de changement « à chaud » de composants défectueux peut permettre dans certains cas une continuité d'activité sans coupure visible pour l'utilisateur. Au-delà, durant la période de Work Recovery Time (WRT), « on observe une superposition de procédures normales et d'opérations manuelles » [22] pour atteindre (repère D du graphique) une situation normale. D'un point de vue de l'utilisateur le Maximum Tolérable Downtime (MDT) représente ses attentes et recouvre la période de coupure ainsi que celle dite WRT [22]. Pour satisfaire ce paramètre propre à chaque métier, il faudra chercher à le réduire. En effet, même s'il ne s'agit pas de l'engagement initial pris au CH du Blanc plus calqué sur les notions de RTO/RPO, cela conditionne la satisfaction des utilisateurs et dirigeants et reste dans tous les cas indissociable : on observe généralement que plus le temps de récupération cible est important (RTO), plus le volume de données à ressaisir est important et donc plus le Work Recovery Time l'est

aussi. « Réduire la durée maximale d'indisponibilité tolérable (MTD) demandera donc d'abaisser le RTO […] ainsi que de diminuer le temps de récupération du travail (WRT) » [22]. Ce dernier paramètre dépend évidemment de l'efficacité des travaux réalisés à la main (ressaisie) mais RTO et RPO sont des paramètres techniques sur lesquels il est possible d'agir et dont l'action aura une influence directe sur le WRT.

2.1.1.3 Mesure de la fiabilité des composants

Finalement, d'autres grandeurs caractéristiques vont nous permettre de mesurer la fiabilité des éléments du système donc du système dans son ensemble. Sans rentrer dans le détail de ces grandeurs notamment en laissant de côté les mesures liées au concept de

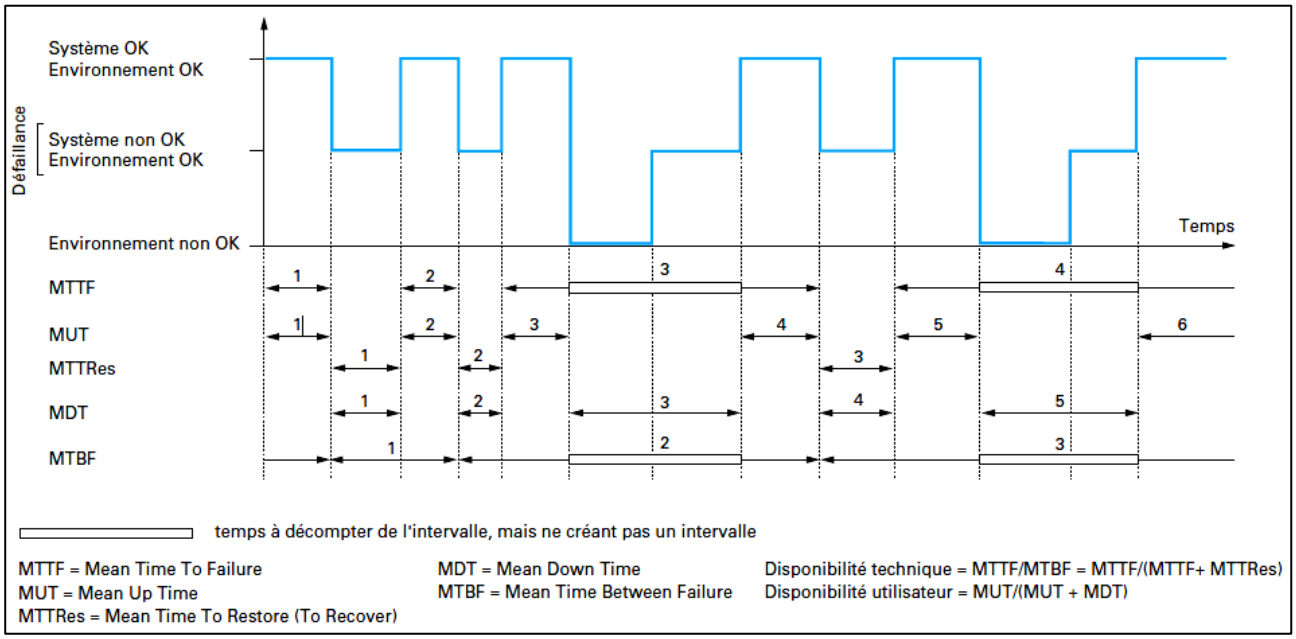


Illustration 16. Intervalles à prendre en compte pour le calcul des mesures de disponibilité [14]

service [14] (le Temps moyen de service rendu (Mean Up Time – MUT) ou non rendu (Mean Down Time – MDT)), ou de maintenabilité (Temps moyen jusqu'à réparation MTTRes ou Temps jusqu'à réparation d'un élément MTTRep), il semble intéressant d'aborder brièvement les « mesures liées au concept de défaillance » [14] dont les valeurs peuvent être communiquées par les constructeurs de matériel (voir graphique ci-dessus).

On rencontre souvent les notions de MTTF (Mean Time To Failure) et MTBF (Mean Time Between Failures). La première mesure le temps moyen jusqu'à défaillance (« espérance mathématique de la durée de fonctionnement jusqu'à défaillance » [14]) tandis que la seconde mesure le temps moyen entre deux défaillances (« espérance mathématique de la durée entre 2 défaillances » [14]). La différence réside dans le fait que le temps de restauration n'intervient pas dans l'estimation du MTTF contrairement au MTBF.

Tous ces éléments nous permettent d'identifier les solutions techniques adéquates afin de s'assurer de la tenue des objectifs du projet.

2.1.2 SOLUTIONS TECHNIQUES

2.1.2.1 Généralités

Sachant que l'objectif d'un système à haute disponibilité, qui est un ensemble complexe, est de masquer à l'utilisateur final les défaillances auxquelles il est sujet, cela nécessite « la redondance au niveau du matériel (au moins) » [14]. En effet, « un modèle redondant permet d'améliorer la disponibilité en multipliant tous ses éléments vitaux par deux. Ainsi, il faudra subir deux pannes au lieu d'une pour rendre le modèle redondant indisponible. » [22]. On parle alors de « tolérance aux pannes » ou « fault tolerance ».

Pour cela, un système doit être vu comme un ensemble décomposé en modules (serveur, un élément d'un serveur, une unité de stockage, etc.) qui deviennent alors des unités de services, de cloisonnement des fautes et donc de réparation [14]. Lorsque l'une des ressources tombe en panne, une autre prend le relais, le temps que le système soit réparé : la disponibilité espérée de l'ensemble se trouve alors grandement améliorée (sous réserve, entre autres, que l'accès entre ces machines ainsi que l'accès des utilisateurs à ces machines soient étudiés avec les mêmes objectifs dans le cadre d'une

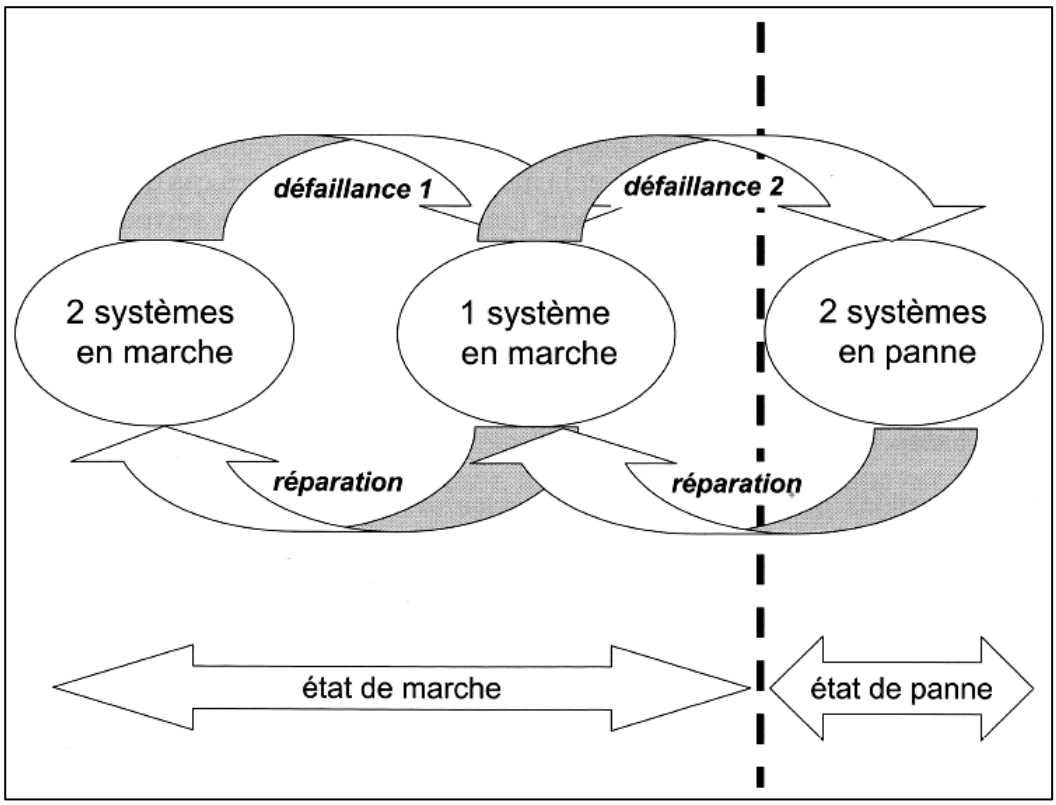


Illustration 17. Fonctionnement du modèle redondant au regard des défaillances [22]

approche globale). Mais de la même manière, le coût de l'infrastructure est lui aussi multiplié…

En outre, « l'installation de modules de rechange (niveau physique) et leur prise en compte dans la configuration du système (niveau logique) doit pouvoir se faire pendant le régime de fonctionnement normal du système » [14]. C'est-à-dire que le système peut, en cas de panne continuer à fonctionner sans rupture de service donc de manière transparente pour l'utilisateur final en se reposant sur le double du « module » défectueux en attendant la remise en état de ce dernier. L'extraction de l'infrastructure du module défaillant, son remplacement ainsi que la reconfiguration en découlant peuvent, dans de nombreux cas, être faites alors que le système reste sous tension : changement d'un disque ou d'un bloc d'alimentation électrique dits hot-plug par exemple. Il s'agit donc d'un système redondant et réparable.

Finalement, il existe différentes solutions basées sur l'ajout de matériel supplémentaire que nous évoquerons plus concrètement ci-après : différents types de modèles dits redondants avec ou sans répartition de charge entre les éléments doublés ou de type « N+1 » [22] [14].

2.1.2.2 Application aux serveurs

Tout comme les éléments qui les composent, les serveurs peuvent aussi être vus comme des modules dont il peut être nécessaire, dans le cadre d'une volonté d'assurer la continuité d'activité, de gérer la défaillance. Différentes publications consultées [22] [14] [24], nous présentent les approches permettant de répondre à ce besoin. Nous savons que tout matériel est sujet à défaillances et que pour pallier celles-ci, il est nécessaire de pourvoir se reposer sur un matériel supplémentaire. Toutefois, différentes solutions existent :

2.1.2.2.1 Le modèle redondant à deux composants

Il s'agit du modèle que l'on imagine aisément à la lecture de ce début de chapitre : la disponibilité est assurée par la multiplication par deux de tous les éléments critiques (alimentation électrique, carte réseau, etc.) donc, ici, des serveurs. Matériel et par conséquent coûts sont multipliés par deux. Le modèle est dit redondant.

Plusieurs variantes existent : « une machine peut être libre pendant que l'autre travaille ou la charge peut être répartie sur les deux en parallèle » [22]. Dans le premier cas, on parle de Failover : en cas de défaillance de la machine dite en production, il y a une commutation automatique sur la machine dite de secours (jusque-là inactive ou occupée à « des tâches indépendantes non critiques » [24]. Il s'agit d'un « recouvrement simple » [24] (Hot Standby ou Simple Failover) dont il existe une variante « recouvrement tournant » [24] où les rôles ne sont pas figés : Lorsque le système anciennement primaire

redevient opérationnel, il se comporte alors comme un système secondaire [24]. Ce dernier cas peut être à éviter si la machine secondaire ou son accès présentent des performances moindres.

Dans le second cas, on parle alors de répartition ou partage de charge ou encore load-balancing. Deux machines se répartissent les traitements tout en accédant aux mêmes données. Attention alors à la prise en compte de l'élément répartiteur de charge qui de fait introduit une nouvelle cause de panne : sa fiabilité doit être prise en compte [22].

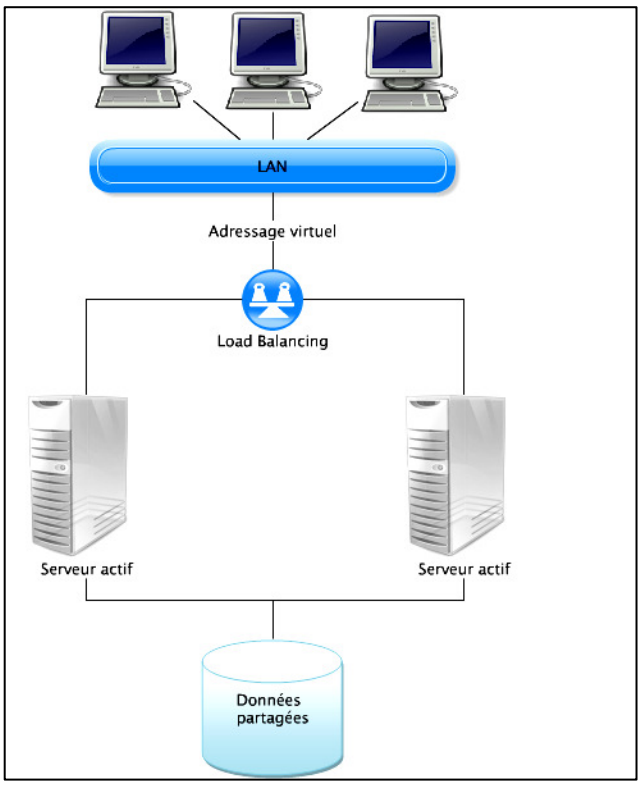


Illustration 18. Modèle redondant à répartition de charge (Load Balancing)

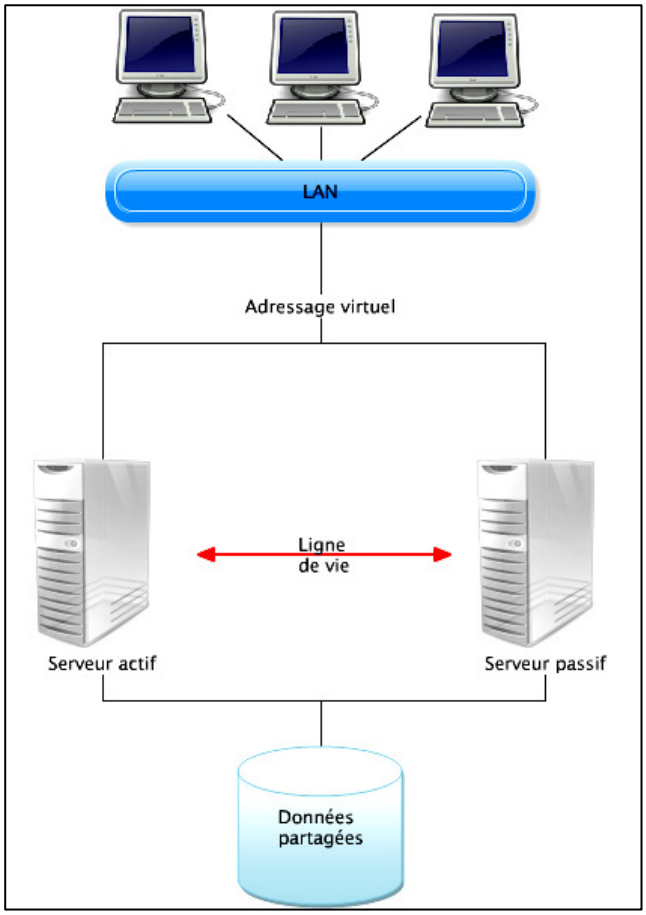


Illustration 19. Modèle redondant de type Failover

Finalement, au-delà de la redondance de serveurs avec ou sans répartition de charges, d'autres répartitions de la charge applicatives sont possibles. Elles sont présentées sur l'illustration suivante :

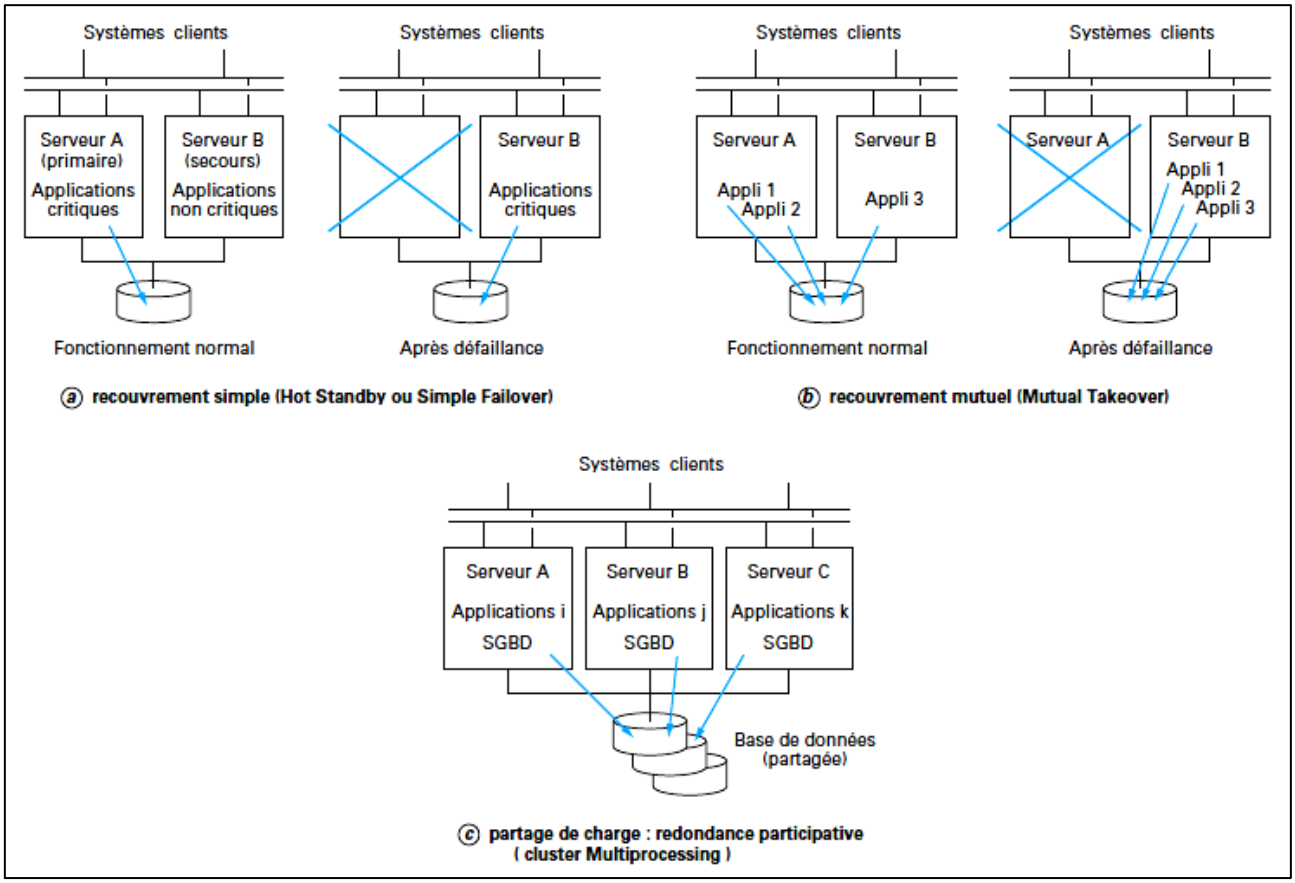


Illustration 20. Différentes approches de fonctionnement en haute disponibilité [24]

2.1.2.2.2 Modèle n+1

« Si N éléments d'un certain type […] sont nécessaires au fonctionnement d'un système, la configuration comporte N+1 éléments » [14]. La charge de travail est répartie sur N machines » [22]. Si l'un des éléments, ici un des serveurs, vient à ne plus fonctionner, l'élément supplémentaire prend sa place jusqu'à ce que l'élément défaillant soit remplacé. Une fois cette étape réalisée, le système est de nouveau prêt à supporter une défaillance de ce type. Dans le cadre de serveurs informatiques, on parle souvent de « cluster ou "grappe" n+1 » [22].

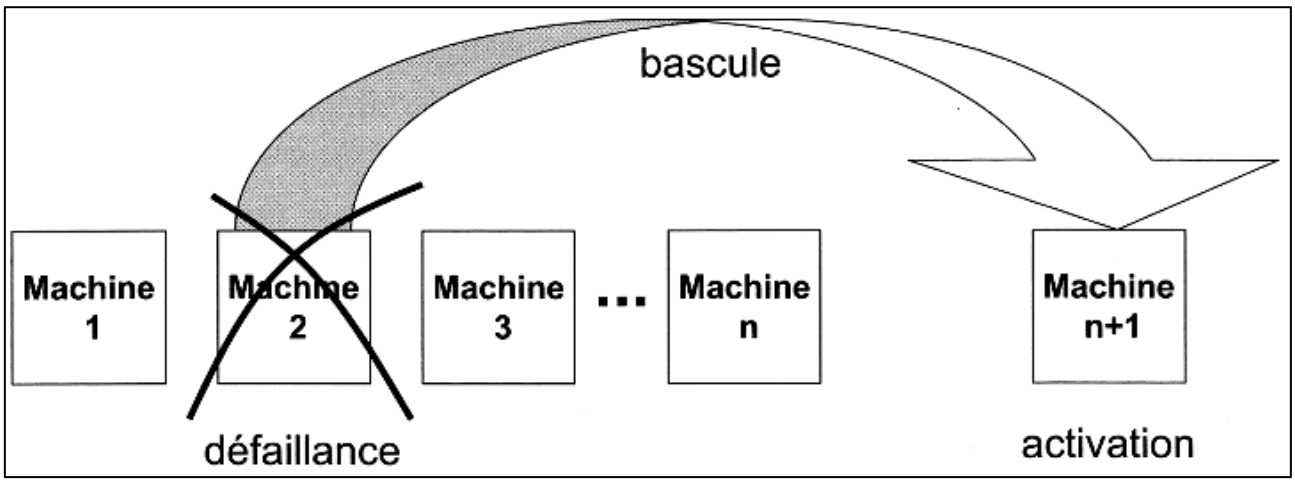


Illustration 21. Représentation du concept n+1 [22]

L'intérêt réside dans la possibilité d'obtenir une infrastructure fiable et performante par la mise en commun des ressources de machines moins puissantes, moins sécurisées et donc moins coûteuses. De plus, la conséquence d'une panne n'affecte que le N^{ème} de la puissance globale donc le N^{ème} de la capacité de traitement. Selon, Emmanuel Besluau [22], c'est la solution souvent plébiscitée par les fournisseurs d'accès à Internet qui choisissent de déployer leur infrastructure sur un grand nombre de serveurs d'un niveau de fiabilité assez commun dont les pannes, facilement réparables, ont des effets peu perceptibles des utilisateurs finaux. La scalabilité (scalability) de l'ensemble de l'infrastructure devient alors possible ou facilitée.

2.1.2.2.3 Surveillance de l'ensemble des constituants du système : Heart Beat

La surveillance de l'ensemble des constituants permettant d'identifier ceux en état de fonctionnement et ceux qui ne le sont pas, est assurée autour de « l'échange périodique de messages entre les constituants du système » [14]. Cet échange est généralement qualifié de battement de cœur ou Heart Beat [14].

2.1.2.2.4 Prise en compte de la panne de mode commun et élimination des SPOF

« Un point de défaillance unique (Single Point Of Failure ou SPOF) est un élément dont la défaillance entraîne la défaillance complète du système » [14]. Il convient donc, lors de la conception d'un système à haute disponibilité, d'éliminer les points de défaillance uniques.

Cela passe par la redondance des matériels mais aussi les chemins permettant l'accès à la ressource [14]. Il peut s'agir, par exemple, de l'élément répartiteur de charge que nous venons de voir, du boitier d'alimentation électrique s'il est unique ou encore le système d'exploitation du serveur. La plupart des points sont facilement éliminés via une architecture adéquate qui, comme vu précédemment peut conduire à la redondance complète du système pour pallier, par exemple, une défaillance du système d'exploitation…

La panne de mode commun est transverse au problème considéré. On peut ainsi imaginer des serveurs différents mais utilisant les mêmes processeurs reproduisant ainsi une erreur commune ou utilisant un antivirus défaillant bloquant l'ensemble des machines. Il en est de même dans le cas de l'utilisation d'une source électrique commune et non secourue, etc. « L'analyse des risques doit donc absolument rechercher ce type de panne générale » [22] afin d'en réduire la probabilité d'apparition ou « en diviser les effets pour éviter qu'elle soit commune » [22]. Ainsi, répartir les traitements et les données dans différents locaux peut permettre de répondre aux effets d'une coupure d'alimentation ou d'un incendie dans ce qui aurait été le local informatique principal… Il est alors nécessaire de s'assurer que les données sont accessibles par l'ensemble des serveurs et que ces derniers sont accessibles de tous les utilisateurs.

En conclusion, il ne sert à rien de mettre en place des éléments en haute disponibilité si la prise en compte de l'environnement d'exécution et les pannes de mode commun ne sont pas prises en compte.

2.1.2.3 Application au stockage des données

Encore plus que pour l'aspect traitement de l'information, le stockage des données n'échappe pas au besoin de disponibilité et plus simplement de sécurisation. De par les

aspects mécaniques des disques magnétiques, la probabilité de défaillance n'en est que plus importante et les conséquences au moins aussi importantes. Les constructeurs ont donc développé différentes solutions permettant la sécurisation et la disponibilité des données.

2.1.2.3.1 Le RAID

Le RAID, proposé dès 1988 par les chercheurs de l'université américaine de Berkley [24] [25] peut permettre de répondre à différentes problématiques dont essentiellement la disponibilité des données (mais pas uniquement : nous verrons, par exemple, la notion de performance, via des données réparties et des accès parallèles, connue sous les termes de data stripping). René J Chevance synthétise l'approche liée à la disponibilité de la manière suivante : « Typiquement, un sous-système RAID dispose de plus de disques qu'il est strictement nécessaire pour le support de la fonction RAID. Les disques en surnombre servent à compléter la redondance : en cas de défaillance d'un disque, les mécanismes RAID assurent la disponibilité de l'information » [24]. Encore, « Un RAID est un ensemble de disques, dans lequel une partie de la capacité utile a été prélevée pour stocker de l'information redondante, permettant la régénération des données des utilisateurs dans le cas où l'un des disques du réseau viendrait à défaillir » [25]. Cet ensemble de disques (ou disk array [25]), présente un ou plusieurs disques virtuels au systèmes d'exploitations qui, naturellement, n'ont pas la vue de l'architecture physique réelle.

Nous présenterons ici les différentes solutions en nous intéressant essentiellement à l'aspect disponibilité offert par le RAID au travers de la redondance.

Les principaux dispositifs RAID (une partie du RAID 0 au RAID 5 identifiés dans l'étude de l'université de Berkley [25] ainsi que le RAID 6) sont en partie présentés sur l'illustration suivante :

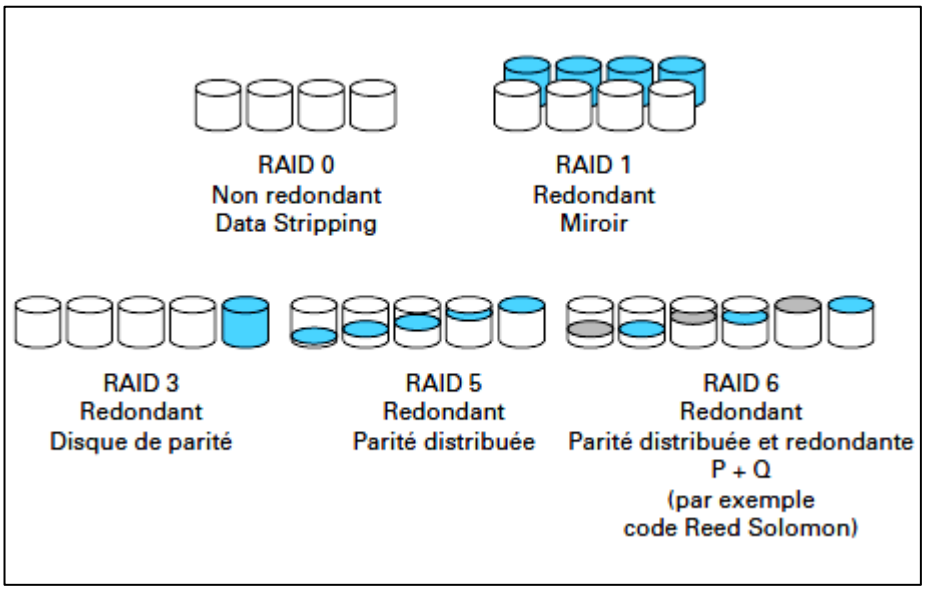


Illustration 22. Principales architectures RAID [24]

- Le RAID 0 dit volume agrégé par bandes : consiste à répartir les données sur plusieurs disques afin de réduire le temps de transfert (lecture / écriture) par des accès parallèles ou présenter un espace disque supérieur. Ce premier niveau n'est que Data Striping et « ne possédant pas la redondance symbolisée par la lettre "R" du mot RAID » [25] donc n'apporte rien en termes de disponibilité des données.

- RAID 1 dit disques en miroir : Le RAID 1 introduit le principe de redondance des données commun à tous les autres RAID via, en ce qui le concerne, la technique du miroir (ou mirroring). Les données étant dupliquées, le principal inconvénient réside dans le coût (utilisation de deux fois plus de disques que nécessaire au stockage des données) mais la lecture peut être faite sur la copie la plus disponible. Cette technique du miroir sera reprise entre les baies de stockage mise en place au CH du Blanc dans le cadre de ce projet de haute disponibilité.

- RAID 3 et RAID 4 : La redondance est alors fondée sur la technique du "ou" exclusif (XOR) [24] permettant ainsi de reconstituer le contenu d'un disque de données ou de parité à partir du contenu des autres disques. Il s'agit dans les deux cas de volumes agrégés par bandes à parité répartie sur un disque supplémentaire (La différence entre ces deux RAID réside dans l'unité de stockage servant de base au calcul de parité). Cette approche est plus performante (les données sont réparties sur des disques synchronisés) et surtout économique (elle ne nécessite qu'un disque de plus que l'espace de stockage nécessaire).

- RAID 5 : Là encore, il s'agit d'un volume agrégé par bandes à parité répartie mais sur l'ensemble des disques. La redondance est là encore économique (comparé au RAID 1) et ses performances sont améliorées par rapport au RAID 3 pour les petits transferts et les mises à jour [24].

- RAID 6 : « similaire au RAID 5 mais complété par une seconde parité distribuée qui lui confère la capacité à résister à deux défaillances » [24].

Ces dispositifs RAID forment ce qu'il convient d'appeler la base. Il en existe toutefois bien d'autres : certains sont obsolètes (RAID 2), d'autres simplement moins répandus ou formés par combinaison des architectures RAID que nous venons de voir (RAID 0+1 ou RAID 10 présentés à l'illustration ci-contre, RAID 50, etc.).

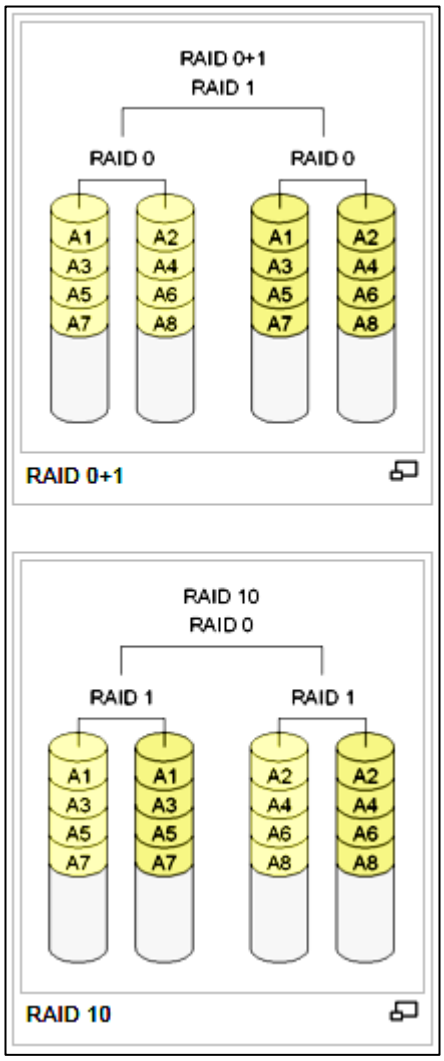


Illustration 23. RAID 0+1 et RAID 10 composés à partir des RAID 0 et RAID 1 [27]

La disponibilité de l'information est assurée par le système RAID qui, après remplacement du disque défaillant (ou des disques dans certains cas), la plupart du temps sans arrêt du système via des échanges « à chaud », reconstitue le contenu du disque nouvellement intégré (avec des délais et des dégradations de performance liés aux matériels et à la priorité donnée à la reconstruction). Au-delà, « Avec des disques disponibles, le sous-système RAID peut décider, suite à une défaillance et tout en assurant la continuité du service grâce à la redondance fournie par l'architecture RAID, de reconstituer la configuration RAID telle qu'elle existait avant cette défaillance en reconstituant le contenu du disque défaillant sur le disque disponible » [24]. La défaillance d'un autre disque est alors permise avant même que le premier disque défaillant ait été changé.

2.1.2.3.2 Dispositifs de stockage dédié et fonctionnalités avancées

« le stockage a pris ses distances vis-à-vis des serveurs » [26]. En effet des dispositifs de stockage en réseau permettent, comme nous allons le voir, de disposer d'une indépendance fonctionnelle axée sur la performance, l'optimisation et la disponibilité des données mais aussi d'une capacité à l'éloignement géographique pour permettre « des opérations de copie de données dans des endroits protégés » [26]. Ce type de dispositif, que nous décrirons à la suite et « indispensable pour bénéficier des aspects de haute disponibilité » [4], intègre dans la grande majorité des cas, la gestion du stockage selon les techniques du RAID précédemment décrites ainsi que l'échange à chaud, au moins, des disques. Cette gestion du RAID est assurée par un ou des contrôleurs. Dans ce second cas, en vue d'assurer la continuité de service, comme pour les serveurs, l'un des contrôleurs peut être en attente (« cold standby » [25]) pour le cas où le contrôleur actif tomberait en panne, ou la charge peut-être répartie entre eux (« dual active » [25]).

D'une manière plus générale, un système de stockage est dit à haute disponibilité car « il est insensible à une panne unique de ses composants » [25] dont : la perte d'un disque,

d'un contrôleur, d'un bus côté disques, etc. De même, « le contenu des mémoires caches est garantie par une batterie de secours, elle-même redondante » [25] et « le système d'alimentation basé sur la technique du N+1 » [25]… Mais comme nous l'avions évoqué concernant les serveurs, « rendre le système de stockage très disponible ne suffit pas, il faut que l'architecture à haute disponibilité s'étende au système hôte » [25].

Sont regroupés sous les sigles NAS (Network Attached Storage) et SAN (Storage Area Network), deux solutions permettant le stockage normalisé des données en réseau. Le premier permet l' « échange de fichiers sur le réseau d'entreprise en protocole de communication de type TCP-IP » [26] pour « des clients évoluant sur des plates-formes et dans des environnements d'exploitation différents » [26] sur différents types de réseaux (Ethernet, Gigabit Ethernet, etc.) avec différents protocoles (NFS, CIFS/SMB, etc.). Son principal attrait est son fonctionnement autonome, ouvert et indépendant des machines auxquelles il offre son service tout en permettant, de par son architecture, une croissance modulaire pour répondre aux évolutions de l'entreprise (scale-out NAS).

Quant à lui, « Le SAN […] est un réseau qui assure la connexion entre des contrôleurs de stockage, des unités de disques diverses et des serveurs » [22] : il s'agit d'un réseau à haute performance dédié aux serveurs et au stockage, non connecté directement au LAN. « Il autorise non seulement les liaisons entre n'importe quel serveur et n'importe quelle unité de stockage mais aussi des liaisons directes entre unités de stockage ou des liaisons serveur à serveur » [26]. On échange, sur un SAN, directement des blocs de données entre machines via le protocole Fibre Channel (FC) voire, grâce à l'essor du réseau Ethernet 10Gb, au protocole iSCSI (appelé aussi SCSI sur IP [22], ou encore FCoE (Fibre Channel over Ethernet)) moins coûteux mais aux performances de plus en plus intéressantes. Cette possibilité d'échange direct entre les baies va permette la

réplication des données d'une baie vers l'autre et ce, de manière synchrone ou asynchrone. Le premier cas est particulièrement intéressant à l'occasion de la mise en place d'une solution en haute disponibilité et le second cas pour la réplication sur un site distant avec lequel les débits ne permettent pas une synchronisation instantanée. Les volumes créés sur les baies de stockage ne sont vus que de manière logique depuis un hôte connecté qui n'y accède qu'au travers de l'adressage d'un LUN (ou unité logique permettant d'identifier un périphérique SCSI). Sans que l'hôte n'ait connaissance de l'architecture technique réelle (il ne « voit » qu'un disque), chaque LUN peut proposer un niveau de RAID voire de réplication baie à baie différent [25].

Apparaissent des baies de stockage qui unifient en leur sein les technologies NAS et SAN. Ces solutions sont connues sous les noms de NUS (Network Unified Storage). Ces trois technologies peuvent également intégrer des supports hybrides conjugués à des fonctions de Tiering (Hiérarchisation automatique des données en vue de l'optimisation de la répartition des données sur les différents supports intégrés).

Tous ces systèmes proposent différentes fonctionnalités permettant d'optimiser les processus de sauvegarde et de reprise : Le snapshot (ou cliché) consiste à « garder une image figée des données pendant un certain temps » [22]. Il s'agit d'une image d'un volume logique, à un instant précis, par enregistrement des modifications apportées au volume logique initial. Cette image peut alors être sauvegardée tandis que la production continue.

2.1.2.3.3 Dispositifs de sauvegarde

Finalement, au-delà des données en production pour lesquelles nous avons abordé la disponibilité voire le mécanisme de snapshot intégré aux baies de stockages en vue d'assurer des sauvegardes, les processus de sauvegarde disposent de leurs propres techniques et outils. Nous les abordons brièvement ici car ces systèmes contribuent à la

continuité de l'activité et à la réduction des temps de reprise. En effet, jusque-là reposant sur des dispositifs de bandes magnétiques pour assurer la sauvegarde des données, les systèmes informatiques, bien qu'utilisant toujours ce support meilleur marché que le disque, se tournent également vers les VTL (Virtual Tape Library). Il s'agit de dispositifs simulant des dispositifs de bandes mais reposant sur des disques (plus performants) afin de réduire la fenêtre de sauvegarde voire par conséquent en modifier la fréquence. Le VTL ne remplace pas forcément le dispositif de bandes mais peut servir d'intermédiaire afin que les systèmes bénéficient de la performance qui le caractérise. Ce dispositif peut contribuer au CDP (Continuous Data Protection / protection continue des données). Le CDP consiste à « surveiller un système en capturant toutes les modifications des données y ayant lieu » [22] et les enregistre sur un système de stockage indépendant, assurant ainsi des points de reprise à partir de n'importe quel moment écoulé que cela soit au niveau bloc ou fichier. On peut alors en imaginer l'optimisation au travers de la déduplication des fichiers ou des blocs : les volumes à dupliquer ou à sauvegarder/restaurer étant moindres, les temps de RTO s'en voient diminués.

2.2 LA VIRTUALISATION

La virtualisation n'est pas un concept nouveau. On peut la définir comme « l'abstraction physique des ressources informatiques » [28]. Bien que l'on parle souvent de virtualisation pour faire référence à la virtualisation de plateformes de type serveur, le concept va bien au-delà. Nous en passerons en revue, à la suite de ce chapitre, les différentes formes en nous arrêtant plus particulièrement sur celles en lien avec les projets d'infrastructure tels que celui présenté au travers de ce document.

2.2.1 LA VIRTUALISATION DE PLATEFORMES SERVEURS : CONCEPTS

2.2.1.1 Définition / présentation

La virtualisation est le fait de présenter à un environnement, une version virtuelle d'un ou plusieurs périphériques existants. On parle ainsi de « disques virtuels, interfaces réseau virtuelles, réseaux locaux virtuels, commutateurs virtuels, processeurs virtuels, et la mémoire virtuelle » [28]. Ainsi, à son niveau et bien qu'il ne soit que rarement cité pour illustrer ce concept, un système d'exploitation a déjà pour rôle de « virtualiser les périphériques et la mémoire d'une machine [dans le but d'] offrir un accès uniforme aux ressources matérielles » [29] pour les diverses applications développées. Ceci ayant pour objectif de masquer l'hétérogénéité en uniformisant l'accès au matériel en profitant de l'abstraction fournie par le système et évitant ainsi le développement d'applications pour chacune des configurations. A un niveau inférieur, certaines cartes intégrées pilotent des périphériques en fournissant une vue abstraite du matériel au système d'exploitation comme cela peut être le cas d'une carte interne RAID faisant abstraction du nombre de disques et tout simplement de la notion de RAID : le système « voit » un disque.

Cependant, de manière générale, lorsque l'on évoque la virtualisation, on fait référence à la virtualisation de plateformes. « Celle-ci correspond à l'utilisation de matériel serveur

pour héberger plusieurs machines virtuelles invitées. Chaque machine virtuelle est un environnement virtuel cohérent sur lequel est installé un système d'exploitation » [28]. Ce concept, très en vogue, n'est pas nouveau puisque déjà présent, il y a environ 40 ans lorsque IBM exécutait des « pseudo-machines » sur ses mainframes. La virtualisation de plateformes existe toutefois sous différentes formes, ci-après décrites.

2.2.1.2 Les différents types de machines virtuelles

On distingue généralement deux familles de machines virtuelles : « les machines virtuelles complètes et les machines paravirtuelles » [29]. La première de ces familles étant elle même divisée en deux : les machines virtuelles concrètes et les machines virtuelles abstraites.

2.2.1.2.1 Machine virtuelle complète

« Une machine virtuelle est une machine complète exécutée de façon logicielle » [29]. Ce type de machine est souvent qualifié de « machine virtuelle » (terme le plus générique). La machine virtuelle complète émule une machine complète y compris son processeur. Dès le moment où il existe une machine virtuelle pour une architecture concrète, il devient possible d'exécuter une application réalisée initialement pour la version réelle donc physique du processeur de la machine virtuelle.

On parle alors d'émulation [28]. L'émulateur n'est au final qu'une application du système hôte. Le système virtualisé n'a pas connaissance de la particularité de son environnement d'exécution : il devient alors possible d'exécuter un système prévu pour un type d'architecture bien précis sur une architecture toute autre. « Le rôle de la machine virtuelle est de transformer du code écrit pour un processeur virtuel en du code pour le processeur concret » [29] (physique).

« Dans la famille des machines virtuelles complètes, on distingue souvent deux sous-familles : les machines virtuelles concrètes et les machines virtuelles abstraites » [29].

Dans la publication, *Virtualisation logicielle : de la machine réelle à la machine virtuelle abstraite*, Bertil Folliot et Gaël Thomas nous précisent qu' « Une machine virtuelle concrète est une machine virtuelle qui exécute du code écrit pour une machine réelle existante » [29]. Des solutions logicielles comme VirtualBox et VirtualPC en sont des exemples. Par opposition, une machine virtuelle abstraite ne possède pas d'équivalent matériel et « fournit un jeu d'instruction complet indépendant de toute réalité physique » [29] et par conséquent non contraint par ces réalités physiques. Les principaux avantages sont de pouvoir disposer d'un environnement d'exécution offrant par exemple des jeux d'instructions plus évolués ou des contrôles plus sophistiqués que ce que proposent les processeurs physiques mais aussi de permettre « d'exécuter du code de façon uniforme sur n'importe quel couple processeur/système d'exploitation » [29]. Parmi les solutions du marché, nous pouvons citer les Java Virtuals Machines, les CLR de Microsoft, etc.

2.2.1.2.2 Machines paravirtuelles

A la différence de la machine virtuelle qui émule l'ensemble d'une machine dont le processeur, la machine paravirtuelle exécute le code directement sur le processeur physique de la machine hôte tout en récupérant une partie (le "code privilégié" : celui qui ne s'exécute qu'en mode superviseur) pour l'émuler [29]. La virtualisation du processeur est donc partielle d'où le terme de « paravirtualisation ». Toutefois, contrairement aux machines virtuelles complètes, les machines paravirtuelles présentent l'inconvénient de ne pouvoir s'exécuter que sur des machines hôtes équipées de processeurs pour lesquels le système invité a été initialement développé. Même si de nombreux systèmes fonctionnent aujourd'hui sur un nombre d'architectures limité, il persiste un lien non négligeable avec le

matériel et, par définition, un niveau d'abstraction moindre, qui peut s'avérer limitant à l'occasion d'un projet de virtualisation. Mais il reste cependant très intéressant car l'exécution directe sur le processeur évite la phase de transformation entre processeur virtuel et processeur physique induite par la virtualisation et rend les machines paravirtuelles beaucoup plus performantes. On parlera alors de « "moniteur de machine virtuelle" ou "hyperviseur" » [29]. Citrix Xen ou VMware ESX sont deux des représentants largement déployés dans le monde de l'entreprise.

2.2.1.2.3 Autres types de machines virtuelles

D'autres approches peuvent être considérées comme de la virtualisation telles que « La virtualisation à noyau partagé » [28] (ou "virtualisation de système d'exploitation", ou "virtualisation au niveau système"), qui permet sous UNIX et Linux, par changement de la racine (chroot), de faire fonctionner un programme ou un système complet dans un environnement protégé. Le système croit alors fonctionner avec son propre système de fichiers sur une machine réelle. Les performances et le rapport entre le nombre de machines virtuelles et les capacités physiques de la machine hôte sont alors très élevées. Toutefois, l'ensemble des machines virtuelles doit impérativement être compatible avec le noyau partagé ; ce qui impose une cohérence accrue des machines virtuelles. Cette approche se retrouve sur Oracle Solaris (ex Sun Solaris) ou OpenVZ.

Proche, « La virtualisation au niveau noyau » [28], est une spécificité Linux qui permet d'exécuter plusieurs environnements. « La machine virtuelle utilise son propre noyau unique pour démarrer la machine virtuelle invitée […] indépendamment du noyau de l'hôte » [28]. Linux KVM (Kernel Virtual Machine) peut prendre en charge différents systèmes d'exploitation invités de type Linux ou Windows.

2.2.1.3 Les environnements d'exécution des machines virtuelles

Tout comme il existe deux grandes familles de machines virtuelles, il existe deux types d'environnements d'exécution que sont : la virtualisation reposant sur un système hôte existant et la virtualisation reposant sur un hyperviseur [28].

Le premier cas est dit « système invité / système hôte », « virtualisation classique » ou « virtualisation hébergée » [28]. Il repose sur un système d'exploitation existant sur lequel fonctionnera une application de virtualisation tierce au travers de laquelle il devient possible de faire « tourner » différents systèmes d'exploitation dits invités. Ces derniers utilisent des ressources partagées attribuées par l'hôte. Cette solution peut présenter l'avantage de ne pas demander une gestion particulière de pilotes pour le matériel puisqu'ils sont déjà opérationnels au travers du système hôte. Toutefois, les performances des entrées-sorties disques restent une difficulté avec cette technologie. Cette approche permet autant la virtualisation complète telle que nous l'avons défini précédemment que la paravirtualisation. Dans ce dernier cas, nous parlerons d'un « Hyperviseur de type II » [29]. VMware Server en est, sans doute, l'illustration la plus connue [28][29].

Le second cas est connu sous l'appellation d' « hyperviseur de type I » [29]. Il s'agit d'une approche dite « bare-métal » [28]. Cet environnement s'exécute sur une machine nue (sans système d'exploitation) [28][29] : « l'hyperviseur est installé directement sur le matériel, puis le système d'exploitation est installé ; il est lui-même une machine virtuelle paravirtualisée » [28] désigné "machine virtuelle zéro". Citrix Xen Server et VMWare ESX/ESXi en sont des applications.

2.2.1.4 Intérêts de la virtualisation des serveurs de l'entreprise

Nous avons évoqué les aspects technologiques sur lesquels repose le principe de virtualisation de plateformes serveurs mais il apparaît intéressant de faire un tour d'horizon

du marché en nous arrêtant notamment sur les grandes fonctionnalités qui font l'attrait de cette technologie pour l'entreprise.

- **Réduction des coûts** : jusqu'alors pour isoler les applications les unes des autres ou la charge induite par leur fonctionnement, la solution consistait à multiplier les petits serveurs. Avec une utilisation de 10 à 15% de CPU en moyenne [20], les serveurs étaient largement sous-utilisés [20][18]. Par une consolidation et une rationalisation qui, sans remettre en cause l'isolation des systèmes, permet la réduction du nombre de serveurs et une réduction des coûts de fonctionnement induits : énergie consommée (un serveur sous utilisé ne consomme pas tellement moins qu'à pleine charge), refroidissement nécessaire [20][18] voire la place au sol qui, dans le cas des datacenters ou des installations conséquentes, se révèle aussi être un argument. Au-delà, selon diverses sources commerciales une économie peut-être attendue au niveau des tâches quotidiennes d'administration : « 73 % des PME qui ont déployé la virtualisation ont constaté des améliorations significatives en termes de temps passé sur les tâches administratives de routine » [18]. Finalement, pour les cas, comme au CH du Blanc pour lesquels il s'avérait nécessaire de dupliquer les machines critiques, cette consolidation aura aussi été source d'économie évitant ainsi la duplication de nombreux serveurs.

- **Flexibilité** : Lorsque la virtualisation est combinée avec l'utilisation de plusieurs machines hôtes et un stockage externalisé (ou simulé par la virtualisation du stockage tel que le propose VMWare vSphere Storage Appliance, VSA), elle peut permettre, par exemple, l'ajout à chaud de ressources de stockage attachées aux VM [20] qui permet « une plus grande souplesse pour gérer l'évolution des besoins informatiques » [21]. Un environnement de ce type permet, très rapidement la

création de machines (VM) à partir de modèles et une allocation automatique de ressources de type CPU, mémoire, etc. [20] (provisioning) pour répondre à des montée en charge au sein d'une machine (scalabilité verticale) voire, dans les datacenters notamment, la création/suppression automatique de VM d'après un modèle pour une montée de dimension supérieure (scalabilité horizontale). Plus simplement, il est possible d'autoriser de manière contrôlée, la création de machines par des utilisateurs (développeurs d'applications, etc.) pour des besoins de test par exemple. Finalement, le déplacement à chaud d'une machine virtuelle d'une machine hôte à l'autre permet aussi la réalisation d'opérations de maintenance planifiées sur le matériel ainsi que, de manière automatique, de placer une VM sur la machine physique la plus disponible ou en vue de libérer certaines de ces machines physiques afin de les arrêter pendant les périodes creuses.

- **Disponibilité** : La virtualisation, lorsqu'elle est combinée à l'utilisation de plusieurs serveurs dans le cadre d'une architecture redondante (telle que nous l'avons décrite au début de ce chapitre), et à un stockage externalisé, permet une continuité et une reprise d'activité grâce à la Haute Disponibilité [18] (intégrant la surveillance des hôtes mais aussi des VM), une simplification des opérations de sauvegarde (via éventuellement les snapshots automatiques ponctuels) et restauration (pour un retour simplifié à un état fonctionnel antérieur [17]), voire à la réplication de VM sur un matériel de secours éventuellement hors site (site distant, site d'un prestataire, etc.) [17] et, pourquoi pas, sur un matériel différent puisque le système n'y est plus lié. Nous évoquions au chapitre 1 les travaux de Taneja Group, qui montraient l'adoption par les PME de la virtualisation dans le cadre de la recherche de la disponibilité de leur système d'information [17]. Cette tendance est confirmée par différents observateurs dont SearchDataCenter.fr [19] qui, placent

l'utilisation de la virtualisation dans le cadre de la haute disponibilité dans les datacenters en second objectif, juste derrière la consolidation.

Proposés par les grands acteurs du marché (Microsoft, Citrix, VMware, etc.), des outils permettent une migration simplifiée. Les outils de P2V (Physical to Virtual), V2P (Virtual to Physical) et V2V (Virtual to Virtual) permettent la conversion d'un système existant vers une nouvelle cible. Majoritairement diffusé et utilisé, le P2V assure pour les systèmes avec lesquels il est compatible et avec un taux correct de réussite la virtualisation d'un système existant. Selon le même principe, le V2V assure des migrations d'une architecture à une autre (Citrix Xen vers VMware ESX par exemple) ou encore le V2P permet un retour vers le matériel.

Finalement, au moment de la réalisation de ce projet, l'intérêt de la virtualisation semble être confirmé par les prévisions de déploiement de nombreux observateurs et notamment le cabinet d'analyse Gartner [15].

2.2.2 LA VIRTUALISATION DE MACHINES DE BUREAU ET DE L'ENVIRONNEMENT DE TRAVAIL

La virtualisation de machines de bureau ou des applications constituant l'environnement de travail de l'utilisateur, est une autre forme de virtualisation très répandue en entreprise. Elle consiste à déporter le système d'exploitation et ou simplement les applications de la machine cliente vers un serveur distant. L'utilisateur peut alors y accéder par l'intermédiaire d'un logiciel ou d'un matériel de type client léger, plus ou moins "intelligent" et ne disposant pas forcement de plus de ressources que nécessaire à suivre l'évolution du bureau transmise par le serveur. Les connexions sont établies au travers des protocoles Citrix ICA, etc. [28].

Plus précisément, il existe plusieurs approches permettant de délivrer virtuellement des applications ou un environnement de travail complet à l'utilisateur final :

- La virtualisation de présentation ou d'applications centralisées (ou Server-Based-Computing) consiste à recevoir sur le poste client une image d'une application ou d'un environnement de travail complet, réellement exécutée sur un serveur. L'outil installé sur le poste client se charge de la connexion, de la transmission des actions réalisées par l'utilisateur via son clavier notamment et se charge d'afficher l'image que le serveur lui transmet en retour. Pour une meilleure mobilité, il est possible de ne pas utiliser un client spécifique et d'utiliser les applications au travers d'un navigateur Web. Ce type de virtualisation présente différents avantages dont la mutualisation et la centralisation des ressources nécessaires à l'exécution des applications. Par conséquent, il devient à la fois possible de sécuriser les machines qui distribuent les applications tout en permettant un fonctionnement reposant sur des machines clientes peu performantes donc peu coûteuses. Il devient ainsi possible de fiabiliser l'environnement de travail tout en en réduisant le coût de possession (TCO) et en offrant un début de mobilité à l'utilisateur. Finalement, les maintenances applicatives s'en voient réduites et fiabilisées. Cette technologie impose toutefois l'utilisation de serveurs performants, peut engendrer des coûts de licence spécifiques et nécessite surtout une connexion permanente, voire performante pour la vidéo notamment, entre le client et le serveur. Malgré les nombreuses évolutions récentes des réseaux opérateurs, la nécessité de cette connexion permanente peut présenter un obstacle à son utilisation pour une population nomade.

  Parmi les principaux acteurs du marché se distinguent Microsoft (Terminal Services), Citrix (XenApp), Systancia (ApliDis), etc.

- La virtualisation d'applications par isolation : consiste à installer (souvent à distance et en streaming) sur le poste client une application qui aura été précédemment packagée sur le serveur avec l'environnement (paramétrage, fichiers système requis, etc.) nécessaire à son exécution. Cette dernière se fera sur le poste client, exploitant ainsi ses ressources, mais dans un environnement cloisonné, indépendant du système d'exploitation d'exécution et autres applications exécutées selon le même principe ou réellement installées sur le poste client. Cette solution, tout comme la précédente, facilite grandement l'administration puisque les applications sont totalement contrôlées par le serveur. De plus elle permet, d'une part, d'utiliser les ressources des machines clientes lorsque le parc le permet, libérant ainsi des ressources serveur et, d'autre part, elle répond à la problématique de mobilité puisqu'il n'y a pas nécessité d'une connexion permanente. Autre intérêt, l'isolation de chacune des applications permet de faire fonctionner sur une même machine, deux applications incompatibles entre elles.

  Cette notion de streaming appliquée aux applications peut aussi l'être au système d'exploitation complet (qualifiée d'OS Streaming ou Provisioning à la volée) : le système est alors chargé de façon sélective au démarrage depuis un disque accessible sur le réseau.

- L'architecture VDI (Virtual Desktop Initiative) : dans ce cas, système d'exploitation et applications sont virtualisés permettant de s'affranchir du support. Il est alors possible d'utiliser des applications ou un environnement complet sur une plateforme totalement différente dans la mesure où il existe une version de l'outil de virtualisation compatible avec le périphérique que l'on souhaite utiliser. C'est, sans doute, la forme la plus populaire actuellement. Elle nécessitait toutefois jusque-là une infrastructure conséquente et surtout une connexion permanente entre le client

et le serveur (dit Hosted Virtual Desktop ou VDI en mode connecté). Plus récemment, l'hyperviseur client permet un fonctionnement en mode déconnecté, éventuellement en utilisant plusieurs OS différents tout en centralisant les images sur les serveurs facilitant ainsi les mises à jour.

- D'autres solutions moins courantes sont aussi envisageables comme l'utilisation de PC en lames (1 lame=1 PC). Dans ce cas, il ne s'agit que d'un déport de l'affichage et des périphériques externes d'une machine unique centralisée dans un datacenter vers, par exemple, un client léger.

2.2.3 LA VIRTUALISATION DU RESEAU ET DU STOCKAGE

2.2.3.1 La virtualisation du stockage

Pilier incontournable de l'informatique, le stockage est lui aussi éligible au paradigme de la virtualisation. Cette abstraction que constitue la virtualisation permet une gestion totalement déconnectée des machines utilisatrices (éventuellement elle-même virtualisées) et l'adjonction de mécanismes pilotés par une intelligence embarquée s'attachant à l'optimisation des performances et de la sécurité.

« Chaque client du serveur de stockage a une vision logique des volumes, ce qui lui masque complètement l'implémentation physique et les technologies utilisées » [25]. Or, comme précédemment évoqué, différents mécanismes visant à répliquer ou optimiser les données peuvent être mis en œuvre par l'administrateur. De manière classique, nous pouvons citer les architectures de type RAID (éventuellement pour le data stripping) ou la réplication entre systèmes de manière synchrone ou asynchrone (présentée sous forme d'un unique support logique). Ces mécanismes sont supportés par les baies de type SAN, NAS, les baies dites de stockage unifié NUS, ou même les solutions logicielles applicables

entre des disques internes à différents serveurs. Ces mécanismes classiques largement déployés assurent sécurité et performance des données ainsi que « l'augmentation de capacité […] par addition de disques à l'intérieur d'un système de stockage ou par connexion d'une nouvelle unité de stockage » [26] au fur et à mesure des besoins.

2.2.3.1.1 Thin Provisioning

Tout comme les mécanismes ci-dessus évoqués masquent la réalité de la répartition réelle des données, le Thin Provisioning consiste à ne pas attribuer physiquement la capacité que l'on présente comme une réalité aux systèmes utilisateurs. En effet, la consommation de l'espace disque étant amenée à évoluer au fur et à mesure de l'exploitation par les utilisateurs et administrateurs, il est nécessaire d'anticiper cette évolution et de prévoir une marge raisonnable pour absorber cette évolution. Cette approche, indispensable dans les environnements non virtualisés présente l'inconvénient de mobiliser dès la mise en production un volume de stockage coûteux, bien souvent sous-exploité ou pire, apparaissant sur certaines machines insuffisant avec le temps. Le phénomène se trouve accentué avec la multiplication des serveurs et la difficulté à estimer finement l'évolution des besoins. Le Thin Provisioning apporte une réponse en attribuant réellement l'espace disque à un LUN à l'occasion de la première écriture (mécanisme dit Allocate On Write). La marge nécessaire au fonctionnement devient alors mutualisée entre les LUN donc optimisée. Il est possible d'allouer plus d'espace disque aux LUN qu'il n'y a d'espace total dans la baie. L'administrateur dispose alors d'une plus grande souplesse d'administration et pourra, par exemple, répondre à un besoin complémentaire par l'adjonction de disques au bout de plusieurs mois : les coûts auront baissé, l'investissement sera étalé et surtout un espace disque inutile aura été économisé. Tout comme pour la virtualisation de serveurs, cela peut aussi se traduire pour des datacenter par une économie d'espace au sol et une économie d'énergie.

Toutefois, le Thin Provisioning peut être source de problème de performance dans les phases d'attribution d'espace supplémentaire. De même, il apparaît important d'anticiper une consommation excessive d'une des applications afin qu'elle ne consomme pas l'ensemble des ressources disponibles. Finalement, afin de répondre à des pics ponctuels, le mécanisme dit Zero Page Reclame permet de libérer l'espace précédent attribué à un LUN lorsqu'il n'est plus utilisé. Il retombe alors dans le stockage disponible pour l'ensemble des LUN.

2.2.3.1.2 Tiering

Comme nous venons de le voir, de nombreux avantages découlent de la mutualisation de l'espace disque au sein d'équipements spécialisés. Toutefois, si le dispositif de stockage est unique et ses performances homogènes, il met sur un même plan des applications stratégiques pour l'organisation et « gourmandes » en accès disques avec des applications peu exigeantes en performances, pas forcément très utilisées mais très « gourmandes » en espace disque. Un tel dispositif nécessite alors de trouver le bon compromis entre espace disque, performance et coût. Pour répondre à cette problématique, les constructeurs proposent des stockages hybrides intégrant, par exemple, des disques SATA de grande capacité à un coût abordable mais aux performances limitées, des disques SAS plus performants mais aussi plus onéreux, voire de la mémoire flash sous forme de disques SSD dont les performances en entrées/sorties sont bien supérieures. « Le déplacement des données est lié au concept de virtualisation de stockage [afin] d'adapter les différentes technologies aux profils de données » [26]. Il est alors possible d'imaginer une répartition manuelle des LUN au sein de ces espaces de stockage en fonction des performances attendues comme se ferait le choix d'un stockage propre à un serveur physique. Le Tiering apporte une solution reposant sur le déplacement automatique des données en fonction de leurs fréquences d'utilisation (on

parle alors de classes de services), d'un type de support ou d'une organisation des supports (type de RAID pour les accès simultanés notamment) à un autre. Il en résulte une optimisation des performances et des coûts de stockage mutualisés entre les différents systèmes consommateurs de la ressource.

2.2.3.1.3 Déduplication de données

Essentiellement déployé à l'occasion des processus de sauvegarde, le principe de la déduplication des données se développe au niveau des baies de stockage assurant le stockage primaire des données de l'entreprise. Le principe consiste à réduire l'espace nécessaire au stockage des données par identification des données dupliquées et leur remplacement par un lien ne pointant que sur une source unique et ce, de manière totalement transparente pour le système utilisateur de ce stockage. Ciblant au départ les fichiers, la déduplication s'intéresse de plus en plus au bloc formant des segments de fichiers plus petits donc plus susceptibles d'être dupliqués. La déduplication permet de réduire l'espace disque nécessaire au stockage donc, comme pour la virtualisation de serveurs, de réduire les coûts, l'espace et l'emprunte énergétique. La réplication baie à baie s'en trouve améliorée.

2.2.3.1.4 D'autres formes de virtualisation du stockage

Il existe d'autres formes de virtualisation du stockage. On peut citer les dispositifs de sauvegarde comme la possibilité offerte par les baies de stockage (entre autres) de faire, indépendamment des systèmes d'exploitation des serveurs, des snapshots. Il s'agit de clichés instantanés d'un volume de données qui n'est, en réalité, qu'une enveloppe dans laquelle on ne viendra réellement écrire que les blocs dans leur état initial au moment de la modification de la source.

Ou encore, dans le but d'optimiser les performances des sauvegardes afin d'en réduire l'impact sur la production, « le système de bandes virtuelles intercale entre les bandes

magnétiques et les systèmes hôtes un stockage intermédiaire […] qui émule les bandes magnétiques » [25]. Ce type de matériel virtuel est communément appelé VTL ou Virtual Tape Library (voir page 51 et suivante). Il présente aussi l'avantage de « masquer aux systèmes hôtes les technologies de bandes utilisées » [25] simplifiant ainsi les évolutions liées au matériel.

2.2.3.2 La virtualisation du réseau

La virtualisation du réseau, une autre forme d'abstraction de la réalité du matériel informatique, permet, là encore, de se soustraire aux contraintes de la réalité physique afin d'en optimiser l'utilisation et l'administration. Elle peut toutefois prendre les diverses formes suivantes :

2.2.3.2.1 VLAN et VSAN

Le VLAN (Virtual Local Area Network) permet « de segmenter ou isoler le trafic en différents domaines de diffusion » [28] sans être contraint par l'environnement physique (connexion des machines à différents matériels actifs ou au sein d'un même matériel, etc.). Ce procédé permet de regrouper des machines selon les besoins de l'entreprise via différents critères techniques : adresse MAC, port de connexion physique, etc. Les communications inter-VLAN ne sont possibles que via les règles définies par l'administrateur qui peut aussi décider de prioriser certains VLAN par rapport à d'autres lors de l'accès à certaines ressources (QoS).

Le VSAN est le pendant du VLAN mais appliqué au réseau SAN. Il permet d'optimiser les performances et de maintenir des équipements de stockage « sans pour autant mettre hors ligne l'intégralité du SAN » [28].

2.2.3.2.2 VPN (Virtual Private Network)

Le VPN est un réseau privé virtuel. Il s'agit d'une technique visant à étendre le réseau privé d'entreprise, via le réseau public ou au moins un réseau plus vaste, à une autre entité appartenant ou non à l'entreprise. Il s'agit de relier une machine ou tout ou partie d'un réseau à un autre réseau via un tunnel virtuel où sont cryptées et encapsulées, via des protocoles dédiés, les données transmises. Cette virtualisation du réseau de l'entreprise est notamment déployée pour relier les sites distants au site principal (ou les sites entre eux), pour relier l'entreprise à certains partenaires extérieurs voire pour permettre la connexion distante des employés nomades.

2.2.3.2.3 SDN (Software Defined Network)

Le SDN apparaît comme une forme de virtualisation du réseau basée sur l'abstraction de la couche matérielle et des spécificités inhérentes à chaque constructeur. Un administrateur peut dorénavant au travers de solutions logicielles standard lui offrant une vue unifiée du réseau, administrer simplement ce dernier, malgré une diversité des matériels et de leurs spécifications. OpenFlow, API parmi les plus déployées, est implémentée par de nombreux constructeurs majeurs dont Cisco, HP, IBM, etc. Ses premières spécifications datent seulement de 2011 et permettent la gestion centralisée des routages. D'autres fonctions comme la QoS devraient prochainement pouvoir être gérées selon la même approche.

3. DEFINITION, PILOTAGE ET MISE EN PLACE DE LA SOLUTION

L'objectif poursuivi, comme nous l'évoquions précédemment, est la mise en conformité de l'ensemble de l'infrastructure matérielle et logicielle avec les exigences de sécurité, de performance, de souplesse et de continuité d'activité que requière le dossier patient informatisé d'une structure hospitalière, tout en tenant compte à chaque étape du peu de moyens financiers et humains propres à ce type de structure.

Nous verrons au cours de ce chapitre comment la solution a été définie pour répondre au mieux aux exigences dans cet environnement contraint. Nous aborderons également les deux grandes phases de la mise en place plus formelle de la solution en insistant plus particulièrement sur la mise en application des objectifs stratégiques (consolidation et disponibilité) ainsi que sur les compromis qu'il aura été nécessaire de faire. Puis, pour conclure cette seconde phase, sur les axes d'amélioration identifiés lors de cette conduite de projet et pendant les mois qui ont suivi la mise en place.

3.1 PHASE DE DEFINITION ET D'ACQUISITION DE LA SOLUTION ET ORGANISATION DU PROJET

Le projet, qui a consisté à la mise en place d'une infrastructure à base de serveurs et baies de disques en haute disponibilité, s'est déroulé sur une période relativement importante. L'étude du besoin, la publication du marché public et la mise en place des principales briques techniques nécessaires au DPI ont été réalisées début 2010 pour un démarrage du service pilote en Mai 2010. L'installation s'est totalement terminée en 2012. L'amplitude très importante s'explique par la multitude de projets menés de front et souvent interdépendants, par la priorité donnée au déploiement logiciel du DPI et par conséquent, pour le projet ici évoqué, la priorité donnée aux composants nécessaires au

DPI reléguant ainsi le reste du projet à des périodes plus appropriées. Cette parallélisation des projets voire la mise en place de solutions provisoires, pour le Wifi notamment, découlent de la nécessité pour l'établissement d'assumer son engagement à informatiser le circuit du médicament à auteur de 50% des lits du champ MCO et 24% de l'ensemble des lits du CH afin de répondre aux exigences réglementaires alors contractualisées avec l'Agence Régionale de Santé dont il dépend. Le choix du mode de marché public de type « dialogue compétitif » pour le choix du DPI (procédure longue mais permettant d'y intégrer tous les acteurs majeurs du processus de prise en charge du patient), le délai imposé pour le démarrage, l'ampleur des chantiers et le peu de ressources internes tant sur les aspects informatiques que relevant du métier, n'auront pas été des facteurs facilitateurs.

3.1.1 PLANIFICATION DE LA SOLUTION ET INTEGRATION DANS SON ENVIRONNEMENT

Sont listés au chapitre premier du présent document, les principaux projets connexes. Ils ont été définis après une analyse globale et systémique visant à définir le périmètre impacté et le découpage en sous-projets cohérents. De cette analyse a découlé la nécessaire mesure d'écart (Patrick Besse [16]) entre l'existant et la cible à atteindre puis la définition des solutions, a priori, les plus adaptées au contexte. Toutefois, sans revenir sur le détail de ces différents projets, le diagramme suivant (sur la base d'un diagramme de PERT simplifié) met en évidence les liens d'interdépendance qui les relient et, par la même occasion, la forte charge concentrée sur quelques mois. De celle-ci a découlé la nécessité de ne mettre en place, dans un premier temps, que les éléments utiles au démarrage. La finalisation de chacun des projets s'est alors poursuivie au-delà du démarrage de l'informatisation du circuit du médicament sur le premier service et en parallèle du déploiement logiciel ; ce qui explique également l'important décalage entre le début et la fin du projet ici évoqué.

Parmi les liens les plus stratégiques d'interdépendance, sources de risques de non tenue des délais, on peut à minima citer :

- Le choix de la solution logicielle pour le DPI et par conséquent de la plateforme d'exécution nécessaire.

- La partie câblage qui est devenue stratégique à court terme pour le LAN (boucle reliant SR1, SR2 et SR3 comme le montre l'Illustration 11 page 16, la remise en état de SR1, le câblage pour les bornes Wifi, etc.) ainsi que pour le réseau SAN (liens nécessaires entre SR1 et SR2) qui allaient être mis en place.

- Le marché de matériel actif qui avait un impact pour le nombre de connexions physiques possibles (bloquant avant même l'ajout de bornes Wifi), pour la sécurité apportée par l'architecture imaginée, mais aussi par l'uniformisation sur l'ensemble de la structure et le découpage en VLAN. Ce dernier permettant de faire cohabiter, tout en maintenant l'isolation, le LAN, les réseaux médicaux (interconnexions et télémaintenances) et, bien entendu, le Wifi dédié au dossier patient.

- Les marchés de fourniture en lien avec les futurs postes de travail. En effet, la mise en place du dossier patient informatisé était donc directement dépendante du réseau filaire, du réseau Wifi, des serveurs virtuels installés sur la plateforme objet de ce mémoire (serveur métier mais aussi serveur de présentation, etc.), du câblage supportant l'ensemble mais aussi des marchés de clients légers et de chariots pour l'exploitation logicielle dans les services.

Par conséquent, une des premières étapes de prise en compte de ce risque de non tenue des délais, au-delà de l'organisation globale et de l'anticipation nécessaire, aura été la

Illustration 24. Vue d'ensemble simplifiée du projet (diagramme de PERT simplifié)

contractualisation avec les différentes maîtrises d'œuvre (MOE) de leurs engagements sur la tenue de chacun des délais.

Le planning reprenant les différents projets liés et sur la période couvrant cette première phase de mise en place aura été le suivant :

Tableau VI : Planning des projets durant la phase 1 de mise en place

Echéances	Actions
27/10/2009	CABLAGE : Publication du marché
30/10/2009	MATERIEL ACTIF : Publication du marché
27/11/2009	CABLAGE : Notification du marché au titulaire
Fin 11/2009	DPI : Choix et notification du marché au titulaire
18/01/2010	SERVEUR & STOCKAGE : Publication du Marché de virtualisation en haute disponibilité des serveurs et du stockage.
19/01/2010	MATERIEL ACTIF : Notification du marché au titulaire
16/02/2010	SERVEUR & STOCKAGE : Notification du marché au titulaire
23 au 25/02/2010	DPI : Audit des processus et formation à l'administration (droits, etc.)
Du 15 au 26/02/2010	MATERIEL ACTIF : Installation « à plat » du matériel.
Mars 2010	SERVEUR & STOCKAGE : Première phase d'installation
02/03/2010	WIFI : Publication du marché.
30 et 31/03/2010	DPI : Transfert des applicatifs de l'ancien serveur vers la VM et préparation pour le DPI
01/04/2010	DPI : Finalisation de l'installation pour DPI (serveur DPI le 1er avril 2010
08/04/2010	SERVEUR & STOCKAGE : PtoV de l'EAI, installation de la machine de management et des applicatifs de supervision.
06/04/2010	DPI : installation des applications clientes sur serveur TSE virtuel.
15/04/2010	WIFI : Etude de couverture sur le pavillon SSR.
19/04/2010	CABLAGE : Câblages SSR terminé.

21/04/2010	CABLAGE : recette des nouvelles fibres optiques et du câblage revu en SSR.
22/04/2010	MATERIEL ACTIF : formation des administrateurs à l'exploitation.
27 & 28/04/2010	WIFI : Finalisation de l'étude de couverture pour le reste du CH.
28/04/2010	VIDAL : Installation des WebServices Vidal (base médicamenteuse) sur une VM.
3 au 5/05/2010	DPI : Formation pharmacie et SSR du 3 au 5 mai 2010
05/05/2010	WIFI : Remise des plan d'implantation des bornes Wifi pour l'ensemble du CH.
19/05/2010	DPI : Démarrage du circuit du médicament en SSR.

Au-delà de cette étape stratégique du démarrage que nous allons voir pour lequel l'ensemble de l'infrastructure se devait d'être pleinement opérationnel et après le délai nécessaire à la difficile stabilisation logicielle :

- Le dossier patient s'est étendu géographiquement et fonctionnellement. La prescription médicamenteuse a atteint, fin 2013, 54% des lits du champ MCO dépassant légèrement l'objectif fixé pour ce premier pallier.

- La mise en place de la solution wifi définitive et l'extension aux autres services ont été réalisées.

- La virtualisation des serveurs physiques existants et l'optimisation de l'ensemble ont été réalisés comme nous allons le décrire par la suite.

- Le firewall, dont le lien direct avec le sujet traité n'a jusque-là pas été mis en avant, a permis l'extension de la prescription médicamenteuse sur l'un des sites distants ainsi que l'interconnexion sécurisée avec le nouveau laboratoire en charge des analyses de biologie.

- Le marché de réfection et extension du câblage informatique a suivi son cours.

3.1.2 DEFINITION DE L'ARCHITECTURE GENERALE ET DU CAHIER DES CHARGES

Au terme de la procédure de dialogue compétitif, la solution logicielle retenue pour l'informatisation du dossier patient était connue et, par voie de conséquence, la liste des prérequis techniques à laquelle notre solution devait répondre : système d'exploitation supporté par l'éditeur, nombre de serveurs nécessaires, SGBD utilisé, ressources consommées côté serveurs et clients, etc. Pour le reste, nous avions précédemment évalué les besoins connexes (voir « 1.3.2 Définition du projet » page 24) et, au travers de l'étude d'impact, les éléments stratégiques à prendre en compte, ainsi que les éléments indépendants que nous souhaitions voir intégrés dans ce marché public (changement du contrôleur de domaine, etc.). Au-delà de la définition de ceux-ci (et de leur évolution prévisible), m'a été confiée la gestion de ce marché public dont, en premier lieu, la rédaction du CCTP dans le respect de la libre concurrence tant pour les intégrateurs soumissionnaires que pour les constructeurs ou éditeurs sur lesquels reposaient les solutions proposées.

Dans cet esprit mais aussi comme le préconise Patrice Selosse [16], consultant et intervenant CNAM à l'UE ENG110 en 2011, afin de laisser les maîtrises d'œuvres potentielles exprimer pleinement leurs compétences, l'expression du besoin formalisée au travers du cahier des charges, se voulait claire dans les objectifs et, en même temps, peu précise dans les solutions en évitant, comme le précise Alain Le Put [16], la sur-spécification et la sous-spécification. De manière synthétique et simplifiée, pour la fourniture de matériels, logiciels et services associés, le cadre imposé [30] était le suivant :

- Solution compatible avec l'environnement logiciel choisi pour le DPI, validé par l'éditeur de ce dernier et reposant uniquement sur des standards du marché pour une meilleure intégration et surtout une meilleure maintenabilité.

- Solution de virtualisation de serveurs en haute disponibilité répartie sur deux salles distinctes (appartenant à deux zones de feu différentes) avec au moins un serveur physique et un élément de stockage dans chaque salle et, de manière générale, intégrant le principe d'élimination des points de défaillance unique (SPOF) par redondance des composants comme évoqué au chapitre 2. La solution devant par conséquent pallier, sans intervention humaine, la panne de tout ou partie d'un serveur, d'une baie de disques, d'un matériel du réseau dédié au SAN voire l'indisponibilité totale de l'une ou l'autre des salles informatiques hébergeant l'installation.

- Les baies SAN (à réplication synchrone et dans un réseau SAN véritablement dédié pour répondre aux exigences ci-dessus) devaient intégrer des disques 15.000 tours / minute, les fonctions de Thin Provisioning pour une mutualisation intelligente de l'espace de stockage (voir page 64), de RAID, d'échange de disques à chaud, etc.

- Une répartition dynamique de la charge (Load Balancing), permettant de répartir la charge sur les serveurs ainsi que sur les baies de stockage voire sur les éléments qui composent ces derniers (accès réseau notamment).

- La fourniture de nouveaux environnements (Contrôleur de domaine Microsoft / serveur de fichiers, serveur Microsoft Remote Desktop Services pour la virtualisation de l'environnement de travail complet de l'utilisateur du DPI voire uniquement de quelques applications, etc.) et la virtualisation par "PtoV" des machines physiques existantes. Certaines « machines » Microsoft devant assurer la redondance du contrôle de domaine.

- Un système de sauvegarde centralisé permettant la restauration ponctuelle et sélective des données voire celle d'une machine virtuelle complète (système et enveloppe) dont les durées d'arrêts quotidiens de production sont précisées quand ils sont possibles.

- La fourniture d'une solution de protection électrique rackable pour l'ensemble du matériel proposé. Celle-ci se devait de pouvoir communiquer avec les serveurs et les baies SAN et d'être de capacité suffisante pour permettre la bascule sur la seconde salle ou l'arrêt « propre » de l'ensemble de l'installation. Ce système ne devait pas être l'unique source d'alimentation des machines de chacune des salles pour, là encore, ne pas devenir une source de défaillance unique.

Elément stratégique majeur, la société DIS, éditrice de la suite d'applications retenue pour le déploiement du DPI, ne validait à cette époque, que la solution de virtualisation Citrix XenServer 5.x comme environnement d'exécution de la plateforme (système d'exploitation Red Hat Enterprise Linux Server release 5.3, base de données Informix IDS, applications) qu'elle commercialisait packagée et maintenait. Bien que compte-tenu des composants utilisés, la probabilité d'une incompatibilité quelconque était très faible (les annonces ultérieures de l'éditeur l'ont d'ailleurs prouvée), les conséquences en auraient été trop importantes pour faire prendre le risque à l'établissement. D'autant qu'au-delà d'un dysfonctionnement avéré, nombre d'éditeurs se refusent à un quelconque engagement de moyens ou de résultat dans des environnements non préconisés car non parfaitement maîtrisés.

Finalement, l'expression du besoin était complétée par un descriptif de l'infrastructure globale en place et en cours de mise en place mais aussi des serveurs en production avec pour chacun les données nécessaires à la validation de la faisabilité et l'évaluation de

l'environnement nécessaire (modèle de serveurs, systèmes d'exploitation, descriptifs et utilisations moyennes et maximums des ressources de type disque, processeur et mémoire).

Compte-tenu du peu de temps disponible pour la réalisation de ce projet, je n'ai pas jugé utile de mener une étude visant à mesurer l'économie potentielle engendrée par le choix de la virtualisation. D'une part, elle aura été peu significative au regard du nombre de serveurs virtualisés et des matériels ajoutés, mais surtout le gain pour le CH se situait principalement au niveau de la consolidation des serveurs et du stockage, entre autres nécessaires à l'exploitation du DPI, sur une plateforme en haute disponibilité.

3.1.3 ORGANISATION ET JUSTIFICATION DU CHOIX ET PRESENTATION DES SOLUTIONS RETENUES

Après publication de l'appel d'offres par la cellule des marchés du CH, deux candidats ont choisi de participer à la visite sur site obligatoire puis de répondre au marché public concerné. Nous détaillerons plus précisément l'offre retenue mais, de par les partenariats existants avec les constructeurs (les deux étaient notamment « 2009 Preferred Partner Gold HP »), et les solutions mises en avant à ce moment par ces derniers, le choix était restreint et les offres facilement comparables, du point de vue fourniture de matériel et engagement quant au résultat attendu, tout au moins.

Les deux offres répondaient aux critères imposés et reposaient sur la même famille de serveurs (HP DL380 G6), les mêmes baies de disques (HP LeftHand série P4000), un dispositif matériel de sauvegarde équivalent (HP Ultrium 1760), ainsi qu'une comptabilisation des licences équivalente. Il est à noter qu'en 2010, Hewlett-Packard était le premier fabricant de serveurs mis en œuvre pour la virtualisation [19]. Les deux soumissionnaires présentaient une liste de références en lien avec ce type d'infrastructure et des certifications se voulant rassurantes quant à leur capacité respective à assumer ce

type d'installation. Finalement, pour répondre aux exigences de l'éditeur de DPI, les deux intégrateurs ont proposé une solution de virtualisation Citrix XenServer 5.6 et, pour répondre aux exigences de haute disponibilité et de migration « à chaud » des machines virtuelles du CH, ils l'ont proposée dans sa version « Citrix XenServer Enterprise Edition ».

Comme précisé page 57 au paragraphe « 2.2.1.3 Les environnements d'exécution des machines virtuelles », il s'agit d'un hyperviseur de type I qui s'exécute sur une machine nue, dépourvue de système d'exploitation (approche dite « bare-métal » [28]). Citrix reconnait la plupart des systèmes d'exploitation à virtualiser au CH du Blanc (mais pas uniquement) assurant ainsi de meilleures performances via le principe de paravirtualisation (voir le paragraphe dédié aux machines paravirtuelles page 55). Le cas des machines dites non supportées sera traité lors de la seconde phase de mise en place.

Toutefois, bien que les points communs soient nombreux, ces offres se différenciaient de par les aspects suivants :

Tableau VII : Comparatif simplifié des offres d'infrastructure

Principaux points de différence	Détails
Volumétrie totale des disques par baie (≠ du volume utile)	Supérieure dans l'offre du candidat retenu (5,4 TO contre 4,8 TO)
Architecture / nombre de serveurs	4 pour l'offre non retenue et seulement 3 mais plus puissants pour l'offre retenue.
Serveur de sauvegarde et administration	Réutilisation d'un ancien serveur du CH pour l'offre retenue et serveur fourni pour l'offre non retenue.
Outil de sauvegarde	Deux solutions logicielles concurrentes proposées mais seule l'offre non retenue avait la capacité à sauvegarder des VM complètes à chaud.
Réseau SAN	4 switchs pour l'offre non retenue contre 2.
Capacité des onduleurs	6000 VA pour l'offre non retenue contre 3000 VA (en lien avec le nombre de matériels).

Le montant global de l'offre retenue était de 88 000 € TTC soit 36% de moins que l'offre concurrente (139 000 € TTC). L'offre non retenue correspondait également aux besoins et, bien que l'espace disque était moindre et les processeurs des machines moins performants, l'architecture à 4 machines assurait la prise en charge d'un nombre supérieur de VM en cas d'indisponibilité totale de l'une des deux salles et la capacité à sauvegarder ces mêmes machines virtuelles était supérieure. La solution en place permet essentiellement de maintenir en fonctionnement les machines qualifiables de « stratégiques » de par la criticité de leur rôle dans le fonctionnement attendu. Par le plus grand nombre de cœurs de processeur et le nombre de machines virtuelles supportées, l'offre non retenue s'avérait aussi supérieure. La fourniture matérielle et logicielle, et la prestation assurées par l'intégrateur pour cette machine complémentaire, pour celle dédiée au management et à la sauvegarde, pour les deux switchs complémentaires, etc., n'ont pas été sans effet sur le coût annoncé. Le point fort de l'offre retenue aura surtout été l'optimisation recherchée par l'intégrateur pour limiter au maximum les coûts sans pour autant nuire à l'objectif. Cet argumentaire aura été décisif quant à l'offre retenue à l'occasion du déroulement du processus de choix. Le recul nous a, jusque-là, prouvé le bien fondé de ce choix qui, tant au départ que dans la durée, aura permis de répondre aux objectifs fixés (même si le processus de sauvegarde aurait pu être plus proche des attentes initiales).

Finalement, d'un point de vue comptable, le coût global de ce projet ainsi que de ceux directement associés et que nous avons cités par ailleurs, a uniquement été imputé à la mise en conformité pour le dossier patient informatisé. Il aurait été possible de trouver une clef de répartition cohérente pour répartir les coûts sur les services de l'établissement en ayant tiré un profit direct ou indirect.

3.1.4 REPARTITION DES ROLES MOE / MOA ET PHASAGE

3.1.4.1 Répartition des rôles MOE / MOA

Au-delà de ce qui pouvait être défini au CCTP (besoins et architecture générale), la répartition des rôles et responsabilités MOE / MOA aura été assez conventionnelle mais pas uniquement. En tant que Responsable Informatique du CH mais plus particulièrement en tant que chef de projets MOA à l'occasion de cette mise en place, mon rôle aura consisté à piloter les différents intervenants dans le respect de la coordination de l'ensemble des projets ainsi que des composantes coût / délais / qualité. La MOE, quant à elle, au travers de la réponse au cahier des charges a défini les matériels et les associations de matériels ainsi que les paramétrages les plus adaptés pour répondre au besoin. Une fois la solution retenue, la MOE avait en charge sa mise en place avec obligation de résultat.

Toutefois, par intérêt pour cette mise en place mais aussi par nécessité d'être en capacité de l'administrer au quotidien, j'ai participé activement à certaines réalisations et la mise en place globale de la solution. Celle-ci a aussi été l'occasion, au-delà du cahier des charges, de réaliser plusieurs choix stratégiques portés par l'expertise et l'expérience de l'intégrateur et après évaluation des risques associés tels que nous les verrons à la suite de ce document.

3.1.4.2 Phasage

Puisque la notion de délai était une contrainte importante et qu'il s'avérait difficilement possible de mener tous les projets de front et de réaliser celui-ci en une seule étape, il aura été nécessaire de découper le projet pour disposer de tous les éléments indispensables au démarrage du dossier patient tout en dégageant le temps nécessaire à la réalisation des autres projets concomitants. Dans la mesure du possible, il n'était

toutefois pas souhaitable, de procéder à des mises en place non définitives qu'il aurait fallu reprendre par la suite.

Les grands ensembles techniques ont été répartis par phase de la manière suivante :

Tableau VIII : Phasage de la mise en place du projet

Eléments techniques dissociables	Phase 1	Phase 2
Mise en place du matériel redondant et de l'hyperviseur en haute disponibilité	X	
Mise en place des onduleurs	X	
Pilotage de l'arrêt par les onduleurs		X
Installation des nouvelles machines pour le DPI	X	
Virtualisation de l'EAI afin de recycler le serveur en machine de management	X	
Virtualisation des anciennes machines en production		X
Découpage de l'espace de stockage pour le DPI	X	
Découpage de l'espace de stockage pour les autres besoins	X	X
Mise en place des sauvegardes de données pour le DPI	X	
Mise en place des exports manuels de machines virtuelles vers d'autres machines du réseau	X	
Mise en place d'export vers le NAS		X
Automatisation des sauvegardes et exports de machines virtuelles		X

Au-delà-du planning des projets interdépendants vu précédemment, la seconde phase se sera poursuivie jusqu'en 2012.

3.2 PHASE N°1 DE LA MISE EN PLACE : LE DPI OBJECTIF STRATEGIQUE

Les éléments nécessaires à un démarrage dans des conditions acceptables de sécurité et de performance compte-tenu du contexte ayant été définis, nous en aborderons ici les différents points clefs de la mise en place et les choix stratégiques qui ont dû être faits à cette étape du déploiement.

3.2.1 PRESENTATION DE L'ARCHITECTURE GENERALE : REDONDANCE ET ORGANISATION DE LA DISPONIBILITE

Les principales caractéristiques attendues reposent sur les approches et techniques énoncées au chapitre 2 du présent document, elles ont été demandées dans le cahier des charges. On les retrouve matérialisées sur le schéma ci-dessous fourni par la MOE (Illustration 25). C'est par ailleurs l'architecture préconisée par HP dans son guide « Building High Performance High Availability IP Storage Networks with SANiQ » [31] ; tout au moins pour l'interconnexion au sein du réseau SAN des baies de disques et des serveurs.

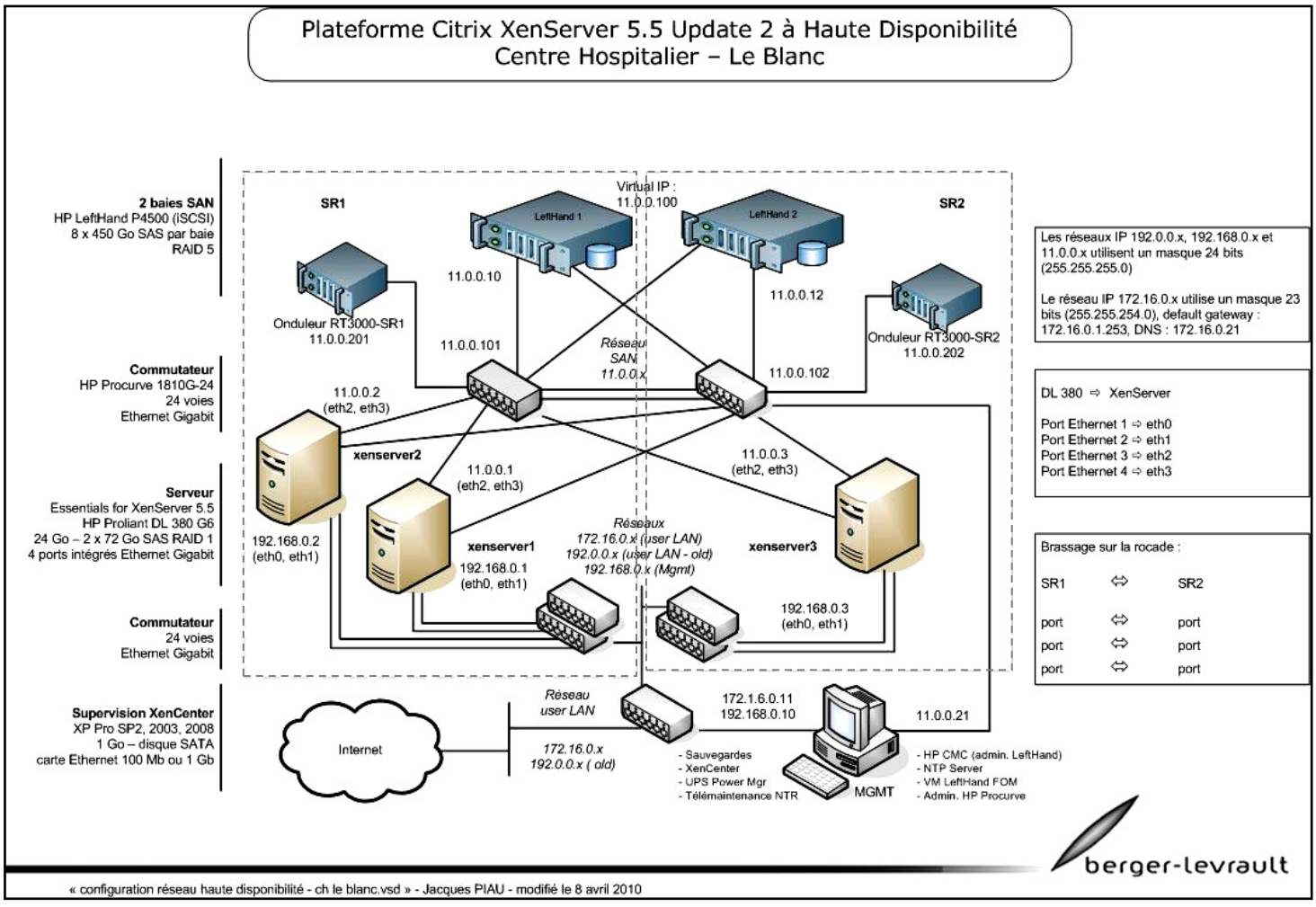


Illustration 25. Représentation générale de l'architecture en place

La répartition des matériels dans les locaux informatiques dits SR1 et SR2 est matérialisée par des pointillés. On constate alors que ces derniers sont tous au moins doublés et que l'on retrouve dans chacune des salles au moins un des éléments assurant une continuité de fonctionnement même en cas de disparition de l'une de ces salles. De manière plus précise, on peut d'ores-et-déjà constater que la redondance est assurée au niveau des baies SAN, des serveurs, des switchs dédiés au SAN. C'est aussi le cas, même si c'est au-delà du périmètre ici traité, pour les switchs dédiés au LAN (au niveau local uniquement).

Bien que cela ne soit pas non plus visible sur l'Illustration 25, les alimentations électriques des machines sont réparties entre la connexion directe au réseau électrique secouru du CH (groupes électrogènes) et la connexion en aval de l'onduleur correspondant. La gestion plus fine viendra, quant à elle, à la phase 2.

3.2.2 SECURISATION DES LIENS D'INTERCONNEXION SAN ET LAN.

A l'exception des onduleurs, tous les différents matériels utilisés disposent d'au moins un accès 1 Gb/s par réseau auquel ils accèdent (LAN, SAN ou les deux parallèlement dans le cas des serveurs). Les réseaux LAN et SAN sont, rappelons-le, physiquement séparés pour des problématiques de sécurité et de performance.

Pour le SAN : les baies de disques constituant l'architecture sont connectées au switch dédié local ainsi qu'au switch dédié distant (de l'autre local informatique) de telle sorte que la panne de l'un de ces matériels réseaux ne bloque pas l'ensemble de l'installation. Cela a été rendu possible par un câblage Ethernet récent et des distances relativement courtes entre les deux bâtiments. C'est toutefois un facteur de risques à prendre en considération puisqu'il n'est pas impossible qu'une forte surcharge électrique sur l'un des locaux (foudre, etc.) se propage à l'autre local via les liens cuivre. Toutefois, au-delà de la probabilité, il

faut apprécier ce risque au regard du dispositif de mise en sécurité de l'ensemble de l'installation électrique dès qu'un risque potentiel d'orage est détecté.
Pour ce qui est des serveurs, leurs connexions au SAN respectent la même logique. Seuls les onduleurs et le futur NAS ne disposeront pas de cette sécurisation.

Pour le LAN : les connexions des serveurs ici évoqués respectent la même approche visant à répartir la connexion (dupliquée) sur deux matériels distincts. Puisque chaque serveur est connecté au matériel actif du local informatique dans le lequel il se trouve, les liens sont répartis sur deux éléments distincts de la pile de switchs qui s'y trouve. Les machines virtuelles hébergées sur ces serveurs ne « voient » qu'une carte réseau virtuelle correspondant à cet agrégat de ports.

La machine dite de management de l'installation ne disposera quant à elle que d'un accès sur le LAN et d'un accès sur le SAN. C'est au travers de ces accès que la vérification quotidienne du bon fonctionnement des équipements réseau est faite.

3.2.3 LE STOCKAGE

Les baies assurent la centralisation et la mutualisation de l'espace de stockage pour faire bénéficier à celui-ci des différentes fonctionnalités évoquées au chapitre 2. Toutefois, toutes celles listées précédemment ne sont pas implémentées dans les baies d'entrée de gamme qui composent notre installation et d'autres n'ont, tout simplement, pas été mises en place.

3.2.3.1 L'organisation de la réplication et de la sécurisation des données

Contrairement à ce que proposent les matériels de plus haut de gamme, nous avons dû définir un niveau de RAID global donc identique sur chacun des volumes que nous allions créer par la suite. Ceci indépendamment du choix de les répliquer ou non d'une baie à

l'autre. Un compromis aura donc été nécessaire pour ne pas réduire l'espace utile par une consommation excessive de disques dédiés à la redondance (ce qui dans certains cas peut permettre de meilleurs taux de transfert mais qui aurait eu un impact à moyen terme sur d'éventuels achats complémentaires et donc le non respect des coûts) tout en assurant un bon niveau disponibilité. Le compromis le plus adapté dans ce contexte et, en même temps celui qui avait été la base d'évaluation initiale, a été de retenir le « RAID 5 » comme organisation de la redondance des données au sein des baies SAN.

Par ailleurs, nous avons défini de manière toute aussi globale, même si ce n'était pas techniquement nécessaire, de répliquer l'ensemble des volumes sur les deux baies. La réplication de type « synchrone » était déjà définie au CCTP, toutefois, nous avions le choix, pour chaque volume créé, de le répliquer ou non sur les deux baies. N'étant pas à ce stade contraints par le manque d'espace disque et surtout animé pas la volonté de rendre, dans la mesure du possible et du raisonnable, toutes les données accessibles même en cas de l'indisponibilité de l'une des salles, il a été décidé que le paramétrage de tous les volumes répondrait à ce critère.

Ainsi organisé, le niveau de RAID réel repose sur le RAID 5 propre à chaque baie complété, via la réplication synchrone inter-baie, par le niveau de RAID 1. Ce niveau de RAID ainsi créé peut être qualifié de RAID 51.

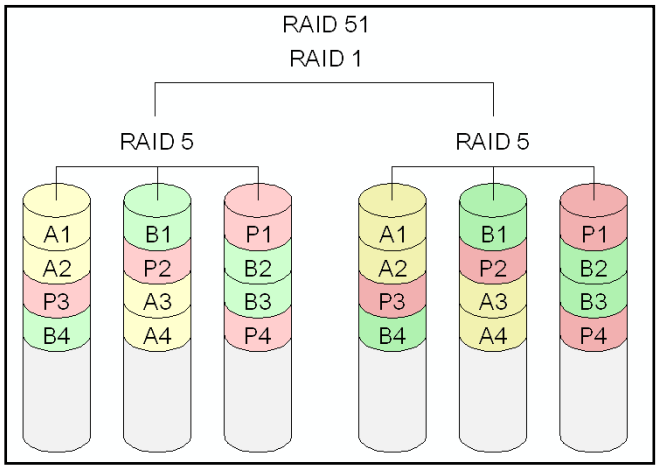


Illustration 26. Présentation du niveau de RAID 51 [27]

Par ailleurs, à ce jour, le niveau de priorisation des échanges liés à la reconstruction du RAID au sein de la baie en cas de changement de l'un des disques, n'est pas paramétré avec un niveau élevé. La question pourra se poser de revenir sur ce point afin favoriser le retour à une situation dite sécurisée au détriment des performances pendant l'opération de reconstruction.

Quant à la répartition de charge entre baies, elle est organisée au travers de l'utilisation d'une IP virtuelle pour l'accès aux données avec l'activation de « Adaptive Load Balancing » [31] .

Finalement, les baies proposées présentent une caractéristique qui n'est pas commune à tous les constructeurs : la nécessité de disposer d'un « Failover Manager » ou encore, un « Gestionnaire de basculement ». Son rôle consiste à assurer le basculement automatique en cas de défaillance de l'une des baies, en jouant le rôle d'une troisième baie chargée d'établir le quorum nécessaire à l'élection de la baie encore accessible. A défaut, en cas de panne d'une baie (voire des liens qui relient ces baies), les données de la baie restante ne sont plus accessibles sans intervention manuelle. La FOM se présente sous la forme d'une machine virtuelle VMware. Dans notre cas, c'est le serveur de management qui l'héberge. Idéalement, il faudrait déporter ce serveur dans un local différent de ceux qui hébergent les baies. Ce n'est, aujourd'hui, pas le cas.

3.2.3.2 Le découpage de l'espace disque

Les serveurs disposent de disques internes de bonne facture en RAID 1 dédiés aux besoins de l'hyperviseur. Aucune des machines que nous avons citées jusque-là n'a vocation à être installée sur ces disques. Les données devant être accessibles de tous les serveurs hôtes et en même temps répliquées. Même si des solutions logicielles permettent d'exploiter les disques internes des serveurs pour proposer des baies SAN virtuelles, ce

n'est pas l'approche choisie. Et donc tous les volumes sont ici créés sur les baies SAN HP (à l'exception, nous le verrons plus tard, des VM pilotant la stratégie de protection électrique).

Le découpage LUN est donc le suivant :

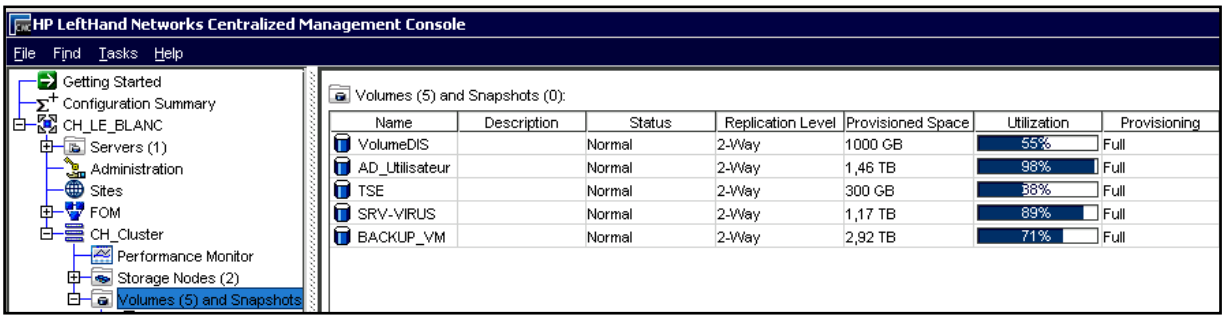


Illustration 27. Vue des volumes créés sur les baies SAN du CH au travers l'outil dédié d'administration

L'utilisation effective est précisée dans le tableau ci-dessous :

Tableau IX : Découpage des volumes

Dénomination du volume	Utilisation
Volume DIS	Dédié à la machine virtuelle Linux, serveur DPI
Volume AD_Utilisateur	Dédié à la machine virtuelle Windows, Nouveau contrôleur de domaine, serveur de fichiers, serveur DNS, etc.
Volume TSE	Dédié à la machine virtuelle Windows, Serveur de présentation pour le DPI.
Volume-SRV-Virus	Mal libellé, ce volume est utilisé pour la virtualisation des disques de tous les autres serveurs de l'installation : Au départ identifié pour la machine devant être virtualisée pour être récupérée pour le management, il a été dédié, lors de la phase suivante, aux autres machines.
Volume BACKUP_VM	Dédié aux copies de sauvegarde des machines virtuelles avant la phase d'export.

Au-delà de la dénomination des LUN qui pourra être revue, ce découpage n'est pas sans poser problème. Nous traiterons cet aspect lorsque nous aborderons la problématique des sauvegardes. Nous aborderons également la non mise en place des fonctionnalités de Thin Provisioning qui, aussi, pose aujourd'hui quelques difficultés.

Sont également stockées sur un de ces volumes (c'est défini lors de l'activation des fonctions de haute disponibilité sur l'hyperviseur), les données liées au Heart Beat nécessaire au contrôle de la disponibilité des serveurs du pool de la plateforme Citrix installée (voir Heart Beat page 44).

3.2.3.3 Les tests et vérifications du bon fonctionnement de l'installation

A ce stade de l'installation, nous utilisons les outils dédiés du constructeur pour valider quasi-quotidiennement (du lundi au vendredi) l'état des baies et des disques. Un contrôle visuel à l'occasion d'interventions dans les locaux concernés et selon une fréquence au moins équivalente, assure une seconde vérification. C'est par ces contrôles qu'est rapidement apparu un dysfonctionnement sur un des disques. Son changement a été réalisé à la suite, selon une procédure classique.

Avant que nous ayons procédé à des tests grandeur réelle, la baie SR2 s'est trouvée partiellement privée d'alimentation électrique au-delà de la capacité de l'onduleur : un disjoncteur de niveau 1 défaillant lors d'un court-circuit a provoqué l'ouverture de celui de niveau 2, commun avec une partie de la salle SR2. Après la coupure brutale de la baie SAN concernée et son redémarrage, nous avons pu constater sa réintégration au nœud et la resynchronisation des données au bout de quelques minutes. Ce test non planifié et relativement brutal, nous a prouvé avant la mise en production, pendant la phase d'installation / paramétrage, que le fonctionnement correspondait à celui attendu (même si l'alimentation électrique était à revoir).

3.2.4 HYPERVISEUR ET SERVEURS PHYSIQUES ASSOCIES

Trois serveurs relativement sécurisés et performants (biprocesseurs 4 cœurs Intel Xeon X5540, 24 GO RAM, disques SAS 72 GO à 15000 tours/minute montés RAID1, double carte d'accès réseau à 2 ports Gigabit Ethernet chacune, double alimentation électrique), répartis sur les deux salles, forment un pool de serveurs dont les ressources sont mises à disposition de l'hyperviseur.

La version de Xen livrée fait de ce pool de serveurs, un pool en haute disponibilité apte, en fonction des paramétrages effectués, à pallier la panne d'une ou plusieurs des machines hôtes par un redémarrage des VM concernées, sur une autre machine du pool. L'arrêt imprévu ou le plantage système d'une VM est aussi surveillé dès lors que les outils Citrix (XenTools) sont installés sur la VM. Partant du principe qu'une salle entière peut devenir inaccessible et compte-tenu de l'offre retenue à ce marché, sans oublier l'étude d'impact réalisée (voir synthèse 1.3.2.2 page 26), le nombre d'hôtes dont la panne serait tolérée a été paramétré à 2 (disparition du local SR1) et les priorités de redémarrage des VM, à ce stade du déploiement, ont été les suivantes :

Tableau X : Priorités de redémarrage des VM en phase 1

Nom et description de la VM	Priorité de redémarrage
Citrix License Server Virtual Appliance	Restart first
Red Hat Enterprise Linux 5.3 (serveur DPI)	Restart
TSE1 2008 R2 (serveur de présentation)	Restart
AD 2008 x64 (Contrôleur Domaine, DNS, etc.)	Restart
Srvenoveai (EAI pour la gestion des interfaces)	Restart if possible

Ce paramétrage sera inévitablement revu lors de la phase suivante et, plus généralement, à chaque ajout de VM ou de changement des caractéristiques d'au moins l'une d'entre elles.

Comme nous l'avons vu, le contrôle du bon fonctionnement des hôtes du pool se fait au travers du Heart Beat qui se matérialise ici via un échange de données stockées sur disques hébergés sur le SAN qui se veut être le lieu d'échange sécurisé et partagé des serveurs.

S'il s'avérait nécessaire d'exploiter des périphériques externes de type lecteur de CD-ROM/DVD-ROM ou périphériques USB : il est bien évidemment possible de faire pointer le matériel virtuel de l'une des VM vers un périphérique de l'un des serveurs. Toutefois, ce matériel attaché à un serveur, ne serait plus joignable par la machine virtuelle si elle venait à être déplacée sur un hôte différent suite à une opération manuelle ou automatique (panne, etc.). Par ailleurs, au travers de la console d'administration installée sur la machine dite de management, nous avons défini, sur cette même machine, une bibliothèque de disques CD/DVD virtuels au travers d'un dossier de fichiers de type ISO accessible de tous les hôtes et des VM qu'ils hébergent. Dans tous les cas, le risque lié à la non disponibilité de ces matériels doit être pris en compte à chaque mise en place.

Là encore, à l'occasion de la panne électrique précédemment évoquée, nous avons eu, au-delà des vérifications quotidiennes de bon fonctionnement via la console d'administration XenCenter, l'occasion de tester en grandeur réelle et de manière là encore brutale, le bon fonctionnement du dispositif en haute disponibilité. Le système ici présenté a parfaitement fonctionné mais a toutefois fait ressortir, qu'au-delà de la disponibilité des machines, il fallait s'intéresser de près aux aspects logiciels impactés par ce type d'arrêt. En effet, lors de cette coupure, le serveur impacté hébergeait le serveur (VM) de présentation dit serveur TSE. Bien qu'il ait parfaitement redémarré, il est apparu nécessaire de gérer de manière automatique au niveau du serveur d'applications et de données du DPI (entre autres), la gestion des sessions plantées et des verrous applicatifs.

A défaut, nombre d'utilisateurs ne pouvaient pas reprendre une activité normale malgré une machine opérationnelle. En collaboration avec l'éditeur, avec lequel je me suis chargé de préciser la définition du besoin, une tâche paramétrée pour s'exécuter très régulièrement élimine depuis les sessions « plantées », et supprime les verrous associés. Ce dispositif a aussi apporté de l'autonomie aux utilisateurs et déchargé le service informatique d'une bonne part de ces opérations nécessaires lors des quelques plantages logiciels ou système côté clients (notamment pour les clients dits lourds généralement utilisés pour d'autres usages que le DPI comme la facturation ou la paie ; et éventuellement en dehors des heures de présence de personnels du service informatique).

Finalement, nous utilisons régulièrement les fonctions de déplacement à chaud de VM d'un hôte à l'autre. Ce fut notamment le cas lors d'un problème sur un des ventilateurs de l'un des serveurs : afin de réaliser l'opération de maintenance, les VM ont été déplacées, le serveur hôte extrait temporairement du pool.

3.2.5 LES MACHINES VIRTUELLES

Différentes machines virtuelles ont été mises en place sur l'hyperviseur selon différentes méthodes que nous verrons à la suite mais toujours avec les ressources préconisées par les éditeurs à l'origine des outils hébergés ou, pour celles concernées, en fonction des ressources réellement consommées sur les machines physiques dont elles étaient issues. L'hyperviseur en place n'a pas la faculté d'allouer dynamiquement (provisioning) l'espace disque réclamé par le système d'exploitation dans la limite d'un éventuel paramétrage. Il est néanmoins possible de revoir à la hausse cet espace de stockage par une opération manuelle. C'est en tenant compte de cette possibilité que les disques virtuels ont été dimensionnés (même lors des phases de PtoV où le redimensionnement est possible). Il

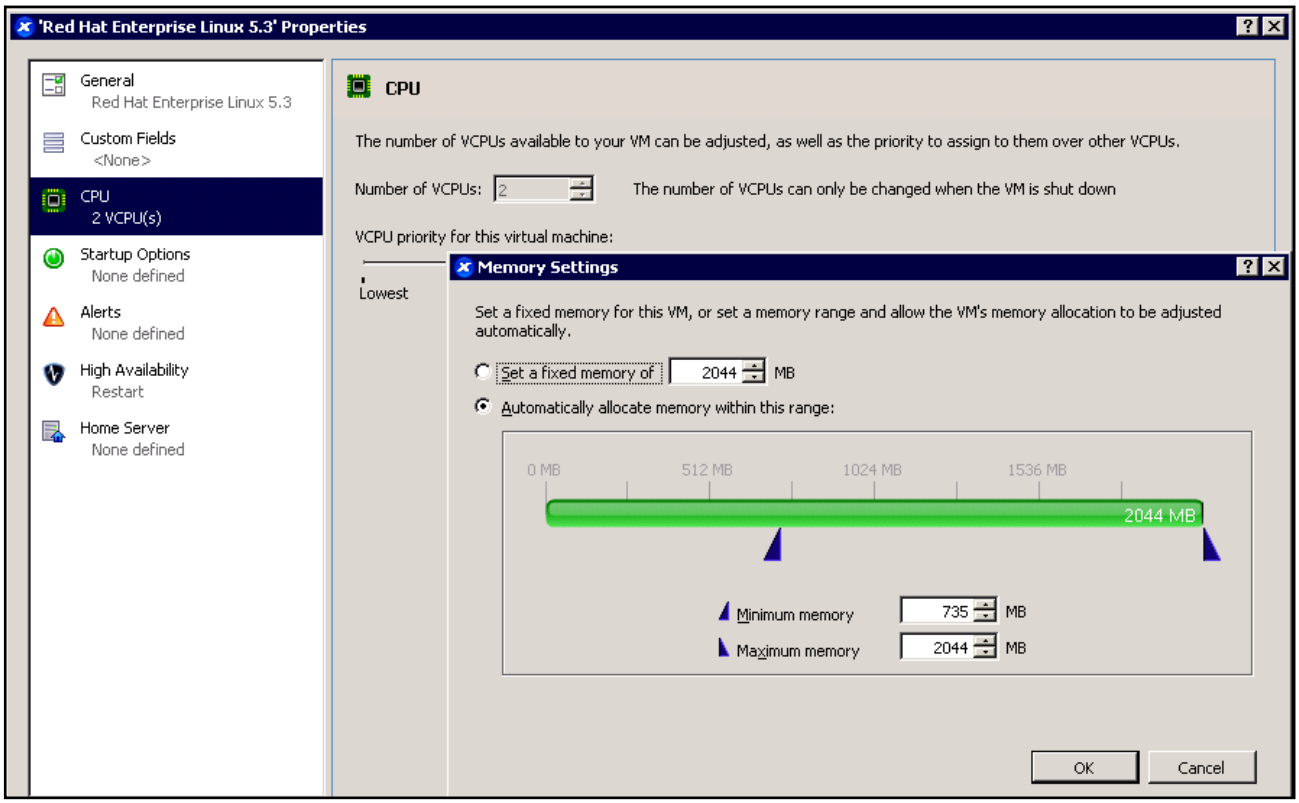


Illustration 28. Exemple de paramétrage des ressources allouées à une VM via l'outil d'administration Citrix XenCenter 5.6

en est de même pour le nombre de CPU où l'opération visant à augmenter ou diminuer le nombre de processeurs reste simple. L'allocation dynamique de la mémoire vive est quant à elle, en partie, activée.

3.2.5.1 Fourniture de nouvelles machines

Pour la fourniture de nouvelles machines virtuelles, l'éditeur de la solution de dossier patient informatisé a fait le choix de fournir une « machine type » standard préinstallée sous forme d'un fichier XVA (XenServer Virtual Appliance) qu'il s'est ensuite chargé d'adapter aux besoins du CH. Au-delà du temps gagné, cette machine présentait l'avantage d'être standard du point de vue de l'éditeur : sans omission que le technicien en charge de l'installation aurait pu commettre et, par définition, sans les inévitables surprises liées aux spécificités du matériel changeant sans cesse. Les recommandations quant aux

ressources ont été respectées et les fichiers correspondant aux disques virtuels sont regroupés dans un LUN spécifique.

Les deux VM Microsoft Windows (le contrôleur de domaine/serveur de fichiers, serveur DNS, etc. d'une part, et le serveur de présentation pour le dossier patient d'autre part), ont été installées selon des méthodes plus classiques consistant à définir le matériel virtuel associé à chacune d'elles, ainsi que le lecteur de CD-ROM (notre cas : simple fichier ISO) qui allait permettre de démarrer l'installation comme si elle se faisait sur une machine physique réelle. Le niveau de disponibilité de ces machines est conforme à ce que nous venons de voir. Toutefois, en complément, pour garantir la disponibilité du domaine Microsoft, par une redondance des contrôleurs, le serveur de présentation est aussi contrôleur de domaine même si, à ce stade, toutes les fonctionnalités ne sont pas répliquées.

3.2.5.2 Virtualisation de l'EAI

Le paramétrage initial de l'ensemble de la plateforme a été réalisé depuis un PC de bureau et, comme nous en avions décidé lors du processus d'achat, il nous fallait récupérer la machine hébergeant l'EAI (machine généralement désignée par SrvEnovEai) pour en faire une machine de management à laquelle serait alors connecté le lecteur de bandes LTO4 externe assurant les sauvegardes de données. Contrairement aux deux premières méthodes, il s'agissait ici de déplacer d'une machine physique, un ensemble système d'exploitation – base de données – applications vers une machine virtuelle. Pour cette première machine, l'opération fut relativement simple et conventionnelle : nous avons utilisé l'outil de conversion de Citrix : XenConvert 2.2.1 dont l'éditeur précise sur l'écran d'accueil « Citrix XenConvert is both a physical-to-virtual (P2V) and a virtual-to-virtual (V2V) conversion tool ». L'utilisation en est très accessible : Lancement de

l'application ; sélection de la source (la machine physique à convertir dans notre cas) et de la destination : XenServer ou, comme nous l'avons fait, via une étape intermédiaire, un XenServer Virtual Hard Disk (VHD). L'étape intermédiaire manuelle n'ajoute pas de risque puisque c'est une décomposition de la méthode automatique. Toutefois, elle aura été utile dans notre cas car le réseau de management vers lequel il aurait fallu faire pointer l'outil de conversion n'était pas visible de l'ancien réseau utilisateur depuis lequel XenConvert était utilisé. Pouvaient aussi être revus à cette étape, les espaces disques : nous les aurons bien souvent réduits : à la fois surdimensionnés par rapport au besoin à l'installation, leur extension est devenue si souple, qu'il est apparu intéressant d'aborder la problématique de cette manière.

Une fois la conversion opérée, il convient de démarrer la machine dans son nouvel environnement matériel : la machine procède alors aux mises à jour de pilotes des matériels nécessaires pour une exécution optimale. Il convient également d'installer les XenTools nécessaires à l'optimisation par la paravirtualisation et nécessaires aux fonctionnalités avancées (déplacement à chaud d'un hôte sur l'autre, arrêt « propre » d'une VM commandé depuis l'hyperviseur, etc.).

La nouvelle machine de management n'avait plus qu'à être configurée pour les nouveaux besoins.

3.2.6 LES SAUVEGARDES

Au-delà de la redondance tant recherchée ici, il n'en est pas moins possible qu'une défaillance matérielle ou logicielle soit la cause d'une perte irrécupérable d'information tout comme, s'il fallait le rappeler, pour les causes d'indisponibilité, la perte de données a bien souvent pour cause, une erreur humaine, ou applicative. Raisons pour lesquelles les stratégies de sauvegarde restent essentielles même pour ce type d'approche. Celle

définie ici reposait sur deux besoins différents à apprécier selon les VM : le besoin de conserver une VM dans un état parfaitement stable d'une part et, d'autre part, de sauvegarder/archiver les données propres à ces machines. Lors de cette première phase, nous n'avons pas encore de NAS opérationnel.

Les objectifs et stratégies prioritaires, à ce stade et par machines, sont :

Tableau XI : Objectifs et stratégies de sauvegarde (phase 1) pour les VM

Machine	Objectifs	Stratégie provisoire en phase 1
Serveur Linux DPI	Conserver une machine stable facilement redémarrable. Conserver des paramétrages/données récents.	Sauvegarde régulière à froid de la VM + exports journaliers de l'ensemble de la base de données en local ainsi que vers un emplacement sur serveur de fichiers…
Serveur TSE	Conserver une machine stable facilement redémarrable avec les dernières versions logicielles. Aucune autre donnée n'y est stockée (fichiers ou paramétrages de l'utilisateur).	Sauvegarde régulière de la machine virtuelle uniquement sans sauvegarde quotidienne de données.
Serveur de fichiers	Conserver une machine stable facilement redémarrable. Conserver des paramétrages/données récents	Sauvegarde régulière de la machine virtuelle et sauvegarde quotidienne de données (dont exports DPI).
Serveur EAI	Conserver une machine stable facilement redémarrable avec les dernières versions logicielles et flux mis en place. Les données récentes peuvent être utiles pour rejouer les scénarii.	Sauvegarde régulière de la machine virtuelle et sauvegarde quotidienne de données.
Machine Management	Conserver une machine stable facilement redémarrable avec les dernières versions logicielles et les paramétrages en place. Les données ne sont pas stratégiques.	Les données sont sauvegardées quotidiennement et une image de la machines est réalisée grâce à l'outil Ghost.

L'organisation des opérations de sauvegarde peut paraître assez conservatrice dans les choix opérés puisque nous de disposions pas, de par l'offre retenue, des agents aptes à sauvegarder les machines virtuelles à chaud et, par ailleurs les snapshots à chaud

exportables réalisés depuis la console de l'hyperviseur (sans même évoquer ceux de plus bas niveau : pilotés pas les baies), ne permettaient pas de garantir la stabilité de l'image sauvegardée notamment au niveau des bases de données embarquées. Deux approches ont été utilisées suivant les possibilités d'arrêt des serveurs. La première consiste à copier manuellement la VM (enveloppe et système de fichiers), après un arrêt, sur un autre LUN dédié aux sauvegardes de machines, avant qu'elle ne soit redémarrée. Cette approche, la plus simple, n'est intéressante que pour les VM dont l'indisponibilité de quelques minutes reste imperceptible ou sans effet. Pour les autres, l'hyperviseur propose la réalisation de « fast clone » que l'on peut apparenter à des snapshots à froid. Sans compter les arrêts et redémarrage, l'opération ne dure qu'une poignée de secondes puisque seule une enveloppe vide est créée dans le même LUN que la VM : après redémarrage de cette dernière, seule l'image originale du bloc de données modifié y est inscrite. C'est alors cette enveloppe qui peut être « copiée » dans le LUN de sauvegarde. Pour ne pas grever les performances globales, il convient de ne pas trop multiplier les « fast clones ».

La technique du Fast Clone n'est pas sans effet sur les stratégies de découpage des volumes sur les baies puisqu'à tout moment de la vie de l'infrastructure, il est nécessaire de disposer au sein du même LUN, d'un espace libre d'une taille au moins égale à la VM à sauvegarder. Sauf à avoir des besoins spécifiques (performance, isolation des espaces disque pour, par exemple, des copies asynchrones sur des sites distants, etc.) ou à disposer d'équipements aptes à réserver et libérer l'espace à la volée, il apparaît économiquement peu concevable de créer un volume pour chaque VM.

Dans tous les cas, pour les machines pour lesquelles un taux de disponibilité élevé est attendu, la sauvegarde à froid ne peut se répéter trop souvent. Elle doit être réservée aux

événements majeurs risquant d'altérer la stabilité du système et doit impérativement être complétée d'une sauvegarde appropriée des données.

Celle-ci, avant d'être affinée, sera quasi-quotidienne : export de base de données quotidien sur une autre machine et sauvegardes sur bande centralisée 5 jours /7.

3.3 PHASE N°2 DE LA MISE EN PLACE : CONSOLIDATION ET PERSPECTIVES

La phase de mise en place initiale a permis de poser les bases de l'installation. Bien que les machines en lien direct avec le DPI aient été opérationnelles et dans un environnement d'exécution raisonnablement sécurisé, de nombreux points restaient à parfaire à l'occasion de cette seconde phase moins contrainte par les délais. Nous traiterons les principaux points dont, la virtualisation des serveurs existants (et la gestion des risques associés) par laquelle nous commencerons puis la haute disponibilité à l'échelle de cette nouvelle installation en abordant à cette occasion la protection de l'infrastructure face aux avaries électriques. Nous verrons finalement comment les processus de sauvegardes ont été optimisés.

3.3.1 LA VIRTUALISATION DES SERVEURS EN PRODUCTION

Après la phase de virtualisation de l'EAI nécessaire à sa transformation en machine de management, les autres serveurs devaient subir la même opération. Le choix en avait été fait dès la définition du périmètre du projet pour pallier essentiellement aux problématiques de performance liées à l'ancienneté des serveurs mais surtout au manque de sécurisation de ceux-ci. Le gain devenait alors considérable de par le dispositif de haute disponibilité en place et de par les extraordinaires possibilités, en particulier de consolidation, offertes par la virtualisation.

Sans revenir sur le détail des opérations, il apparaît intéressant de s'arrêter sur les grandes problématiques rencontrées, les solutions trouvées mais aussi sur les risques qui ont été jugés acceptables ou non dans les prises de décision.

3.3.1.1 Les contraintes de migration

3.3.1.1.1 Licences OEM des serveurs

De par l'approche qui visait à associer un serveur aux besoins d'un éditeur et à la relative stabilité de ces besoins, il devenait économiquement intéressant sans que cela ne devienne un frein à l'évolution ou au maintien, de conserver sur un serveur, le même système d'exploitation tout au long de la vie de la machine. Pour cette raison, l'ensemble des serveurs avec système Microsoft à virtualiser était équipé de licences de type OEM (donc indissociables du matériel avec lequel elles avaient été commercialisées et sans possibilité d'évolution majeure). Comme j'ai eu l'occasion de le préciser dès le processus d'achat, de nouvelles licences ont été fournies et ont dû être réenregistrées sur les systèmes Windows dès le démarrage de ceux-ci dans le nouvel environnement matériel (matériel virtuel). C'était une contrainte contractuelle et contrôlée techniquement pour les versions 2003 et suivantes.

3.3.1.1.2 Les aspects techniques et organisationnels de la migration

Sauf exceptions, les migrations ont été réalisées selon le processus décrit précédemment pour l'ex-EAI. Toutefois, il aura été nécessaire d'organiser les coupures en tenant compte de la durée de l'opération et des possibilités d'arrêt de production de chacune des machines.

Globalement :

Tableau XII : Contraintes et organisation des arrêts pour les opérations de PtoV

Serveur (Nom et rôles)	Contraintes de coupure et choix possibles
SrvQHS : Serveur d'applications dédiées qualité, gestion des risques)	Utilisation continue et importante aux heures de bureau par les gestionnaires des risques et qualiticiens, et utilisation discontinue 24h/24 pour les déclarations d'événements indésirables ou la consultation de procédures qualité. ➔Migration possible dès les fins d'après-midi jusqu'au lendemain matin.
PMSIPilot : Analyse activité hospitalière, Contrôle de gestion, Qualité facturation/codage, etc.	Utilisation ponctuelle sans réelle contrainte sauf analyses périodiques identifiables à l'avance. ➔ Migration possible en journée
Sirilog : Système d'Information de Radiologie (RIS)	Utilisation 24h/24 mais continue et soutenue en journée et, en fonction des urgences, la nuit. ➔ Migration envisageable en fin de journée sous réserve d'une activité peu intense au service des urgences et avec possibilité d'arrêt de la procédure en cas de besoin.
SrvVirus : Gestion centralisée de l'antivirus.	Non indispensable. ➔ Migration possible en journée.
SrvWinrest : Gestion des commandes de repas patients Winrest	Utilisation discontinue mais organisée autour de tâches automatisées planifiées essentiellement le matin et en début d'après-midi. ➔ Migration possible en milieu de matinée ou à partir du milieu d'après-midi.

Certaines migrations comme pour le SrvQHS ont été réalisées en fin de journée car cette organisation n'occasionnait pas de coûts ou autres difficultés supplémentaires (allongement de la prestation, etc.). Si tel n'avait pas été le cas, nous aurions procédé à une coupure en journée.

Juste effleuré au paragraphe « 1.2.2.2 Infrastructure réseau du CH et des sites distants », j'ai recherché de manière la plus transparente possible à conjuguer ces différents projets en vue de tendre progressivement vers la cible définie à l'échelle globale. En effet, il apparaissait nécessaire d'étendre notre plage d'adresses réseau disponibles tout en se

conformant aux recommandations RFC1818 traitant des «Address Allocation for Private Internets » [12]. Cette opération nécessitait, la mise en place des VLAN et du routage de niveau 3 permis par la nouvelle infrastructure réseau (afin d'assurer une migration progressive). Pour plus de commodité, cette opération nécessitait également d'avoir intégré le nouveau paramétrage notamment DNS supporté par une des VM mises en place lors de la première phase de déploiement : il m'était, en effet, apparu préférable de n'interconnecter que le nouveau VLAN utilisateurs (correspondant aux critères ici évoqués) aux accès LAN des nouvelles machines. Ces différents choix imposaient alors de migrer les adressages des serveurs au maximum au moment de l'opération de PtoV. C'est ce que j'ai choisi de faire.

A titre d'exemple, pour la machine Sirilog, la migration de la machine vers l'environnement virtuel et le changement d'IP associé avaient été anticipés par une migration vers le nouveau domaine des machines clientes et la préparation des paramétrages à l'utilisation de la résolution de noms DNS offerte par le nouveau serveur. Une fois le processus de PtoV déroulé et validé, la modification des entrées DNS du serveur terminée, les machines clientes se sont reconnectées comme à l'habitude ; fiabilité et performance en sus.

Le serveur antivirus, quant à lui, touchait tous les PC et serveurs du CH. Il n'était pas envisageable d'attendre la migration de l'ensemble pour procéder à sa virtualisation. Autre difficulté : les PC Portables connectés par intermittence. De par ce fonctionnement un peu particulier (des machines clientes autonomes viennent récupérer auprès du serveur des stratégies et des données voire des versions applicatives), l'ancienne version du serveur et la nouvelle ont cohabité sur une courte période : la stratégie diffusée par l'ancien serveur précisant la nouvelle IP pour les échanges suivants.

3.3.1.2 La gestion du risque lié à la virtualisation

Une opération « classique » de virtualisation selon la technique du PtoV consiste, par une opération logicielle, à faire une image d'une machine physique à transférer directement ou indirectement celle-ci vers un hyperviseur via notamment des formats standards. Toutefois, le fonctionnement de cette opération n'est pas garanti. Dans le cas du serveur dédié aux applicatifs qualité et gestion des risques (SrvQHS), la modification de l'environnement « matériel » vue par le système d'exploitation une fois virtualisé a posé problème puisque le système n'était plus en capacité de démarrer normalement (même après recherche et nettoyage en mode démarrage sans échec, etc.). Une solution existait forcément mais le choix a été fait de ne pas passer trop de temps dans ces recherches qui, même si elles avaient abouti, n'auraient pas été économiquement raisonnables et n'auraient peut-être pas permis de garantir la stabilité du système dans le temps. Etant propriétaire de la licence Windows nécessaire (initialement prévue pour la migration OEM), nous avons préféré commander une réinstallation à l'éditeur (prestation pour un total de 1172 € TTC) et avons choisi l'économie et surtout la sécurité par élimination de la source du risque.

D'autre part, le cas de la machine Sirilog était quant à lui différent. La virtualisation de la machine a parfaitement fonctionné mais l'hyperviseur ne supportait pas le système d'exploitation (Windows 2000) utilisé par cette machine. Ce système n'étant lui-même plus supporté par Microsoft. J'ai jugé raisonnable de continuer la démarche sachant que l'on se retrouvait alors dans un environnement totalement virtualisé et non pas paravirtualisé (les XenTools n'étant pas compatibles). Par conséquent, la VM allait simplement évoluer dans un environnement matériel différent sans possibilité de migration à chaud, etc., que les performances même en virtualisation complète restaient supérieures à celles du vieux serveur physique, et que la sécurisation de l'ensemble au-delà de la compatibilité restait

aussi supérieure. Il est à noter que d'autres établissements contactés faisaient « tourner » des machines Windows 2000 dans un environnement similaire. En cas d'incompatibilité avérée, le serveur physique restait disponible pour faire marche arrière. Le temps a prouvé que cette prise de risque à priori raisonnable et accompagnée d'un plan de levée de risque (Patrice Besse nous précise à ce sujet que l' « on peut prendre des risques mais ils doivent être limités et accompagnés d'un plan de levé de risques » [16]) aura été le bon choix.

Finalement, le serveur PMSIPilot regroupait les deux difficultés ici évoquées : un système d'exploitation non supporté par l'hyperviseur et un PtoV inutilisable puisque les partitions Linux n'étaient plus visibles après opération. La machine n'étant absolument pas stratégique en terme de données (pas de saisie mais import de données standardisées pour analyse), différentes solutions ont été étudiées dont l'utilisation de produits de PtoV concurrents assurant la conversion vers les formats standards ou les outils de sauvegarde Linux mais sans réussite. Le choix a alors été fait de procéder à une réinstallation via la fourniture d'un CD (package système d'exploitation + applications en installation automatique) et ce malgré la non prise en charge de Linux Mandriva par Citrix. La prise de risques restait ici contrôlée et la solution choisie donne, à ce jour, entière satisfaction.

Finalement, les prises de risques auront toujours été contrôlées ou refusées (SrvQHS) tout en tenant compte de la disponibilité des ressources pour les recherches. Toutefois, l'inconnue en était sans nul doute la pérennité notamment en lien avec les évolutions des versions de l'hyperviseur. Cela restait toutefois à relativiser car les machines Sirilog et SrvQHS hébergeaient des applicatifs plutôt en fin de vie et que la VM PMSIPilot, de par son utilisation, laissait le temps à la recherche de solutions. De même, les serveurs physiques supportant ces applicatifs étaient aussi en fin de vie. Au besoin, le retour vers

un environnement non virtualisé n'aurait pas eu beaucoup plus d'impact que le simple changement de serveurs auquel il aurait, dans tous les cas, fallu procéder.

3.3.2 L'ORGANISATION DE LA DISPONIBILITE

3.3.2.1 La gestion de la protection électrique

Cette seconde phase de mise en place aura aussi permis d'aller plus loin dans la configuration d'un système de protection électrique sans remettre en cause les premières étapes de celle-ci. L'objectif initial était double : d'une part, pallier de manière transparente une coupure électrique de courte durée ou une simple variation de tension voire, d'autre part, à protéger l'ensemble de cette installation en en pilotant l'extinction lors de coupures prolongées. La première étape n'avait consisté qu'en la répartition des connexions des matériels sur l'onduleur du SR auquel ils étaient attachés et sur le réseau du CH secouru par les groupes électrogènes. Il restait alors cette étape plus technique de communication entre les éléments constituant la plateforme et le système de protection électrique.

La solution mise en place pour répondre à ce besoin est sans doute artisanale et perfectible mais constitue une première approche intéressante. Elle ne constitue pas un moyen à la continuité d'activité mais elle y contribue par un arrêt propre et contrôlé des VM et des baies de disques laissant l'ensemble dans un état stable prêt à être redémarré dès les ressources nécessaires à nouveau disponibles.

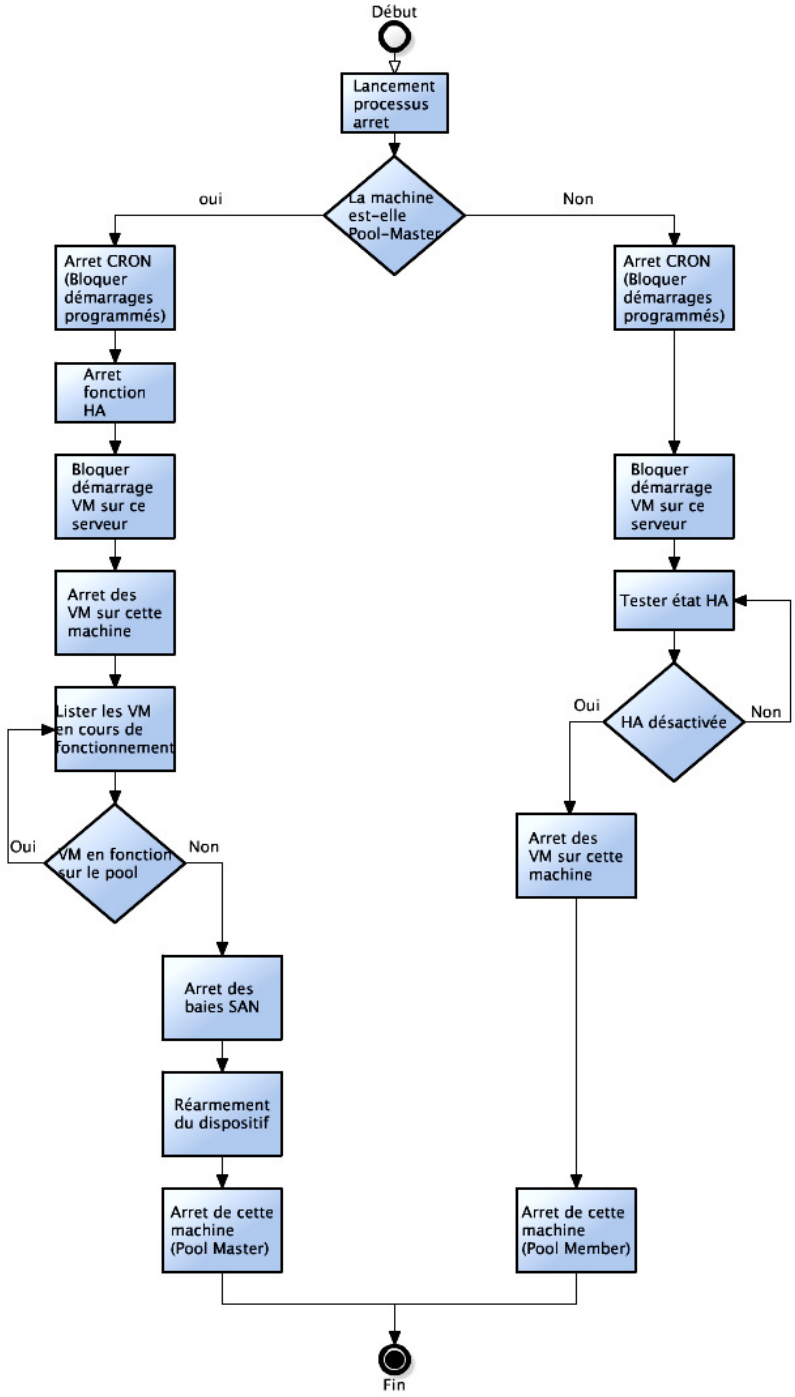
Le principe détaillé à l'Illustration 29 ci-après est le suivant : les onduleurs des SR1 et SR2 sont connectés au réseau IP du SAN (voir Illustration 25 page 83). Une machine virtuelle Linux dédiée, par salle, communique en permanence avec l'onduleur correspondant au travers des agents mis à disposition par le constructeur.

Le processus d'arrêt de l'infrastructure en cas de panne d'alimentation électrique

Sur chaque serveur physique est installée en local une VM fonctionnant sous Linux avec l'agent de contrôle de l'onduleur local installé (une machine par salle ou onduleur doit être démarrée).

En cas de panne électrique de plus de 3 minutes, l'agent exécute un script Linux qui, au travers d'une connexion à un partage CIFS, copie sur la machine de management, un fichier « alerte_onduleur ».

Nb : pour rappel, dans le cadre d'une coupure générale EDF, deux groupes électrogènes en « haute disponibilité » avec répartition de charge en prennent le relais. Ce mécanisme présente essentiellement un intérêt en cas de dysfonctionnement interne.



En amont, une temporisation de quelques secondes permet d'interrompre le déclenchement du processus. Au-delà, une fois lancé, il doit aller à son terme.

NB : il n'existe pas de lien direct entre les VM et les onduleurs puisqu'au-delà de la séparation du LAN et du SAN, les machines ne seraient pas en mesure de rester à l'arrêt puisque protégées le dispositif de Haute Disponibilité de l'Hyperviseur. Toutefois, pour les VM non équipées des XenTools, cette autonomie quant à leur arrêt reste nécessaire mais doit être intégrée dans la procédure globale pour ne pas venir en conflit avec la HA.

Illustration 29. Processus de gestion d'une panne électrique prolongée

En cas de défaut d'alimentation électrique prolongé, la VM exécute un script chargé de déposer dans un emplacement partagé un fichier pilotant l'arrêt des machines constituant l'infrastructure. Indépendamment, sur chaque serveur physique hébergeant l'hyperviseur, une application teste à intervalle régulier l'existence dudit fichier et commande au besoin l'arrêt des machines selon une chronologie précise.

Cette installation telle qu'elle a été mise en place pose le problème du SPOF puisque le partage permettant l'échange d'information est situé sur le serveur de management qui par conséquent devient stratégique. Ce point constitue donc un point d'amélioration.

D'autre part, l'état des onduleurs ainsi que la connexion des agents sont vérifiés quasi-quotidiennement et le remplacement des batteries planifié à intervalle régulier. Le tout est bien entendu couvert par un contrat de maintenance.

Les tests de bon fonctionnement, toujours indispensables ont mis en avant une incohérence entre le fichier d'alerte déposé et celui recherché rendant la solution inopérante. Le problème a bien entendu été traité mais prouve à nouveau, s'il en était nécessaire, le besoin de réaliser ces tests. Le premier niveau, ayant permis d'identifier ce dysfonctionnement, aura consisté à déconnecter l'alimentation électrique en amont des onduleurs, à tour de rôle. Dès lors l'arrêt s'opère au bout du temps imparti. Toutefois, ce test n'était qu'une simulation permettant de s'assurer du bon fonctionnement des automatismes mais elle ne prenait pas en compte le problème dans sa globalité. En effet, lors d'une coupure réelle sur l'ensemble du bâtiment, il s'est avéré que l'échange de messages entre les serveurs supportant l'hyperviseur se faisait via le réseau LAN. Or, ce dernier, n'étant pas électriquement protégé, seul le pool master est resté opérationnel. Les VM protégées par le dispositif HA ont, bien entendu, redémarré mais ce n'était pas le fonctionnement attendu. Depuis, les matériels supportant le LAN ont été protégés mais,

plus globalement, c'est davantage l'approche qui devait être revue : continuer, voire développer les tests sur les composants mais surtout réaliser des tests à l'échelle de l'installation pour s'assurer du bon fonctionnement dans des conditions réelles.

Toutefois, pour revenir à l'objectif, on peut remarquer une limite à la continuité de fonctionnement : la non-distinction des salles. Elle aurait été possible par déplacement des VM importantes sur une machine du pool située dans l'autre salle, en promouvant éventuellement un autre pool master, en découplant les baies pour arrêter celle où l'alimentation pose problème, etc. La recherche d'une telle solution aurait permis une plus grande disponibilité mais nous ne souhaitions pas supprimer temporairement la redondance du stockage pour la prise en compte d'une indisponibilité exceptionnelle et forcément prise en charge par les équipes techniques de garde.

Finalement, pour plus de sécurité, la gestion du redémarrage reste manuelle. Là encore, ce n'est pas la disponibilité qui a été privilégiée mais la cohérence et la stabilité de l'ensemble. En effet, bien qu'il soit possible de temporiser l'alimentation en aval de l'onduleur après retour du courant et bien qu'il soit possible de temporiser le démarrage des serveurs après qu'ils soient de nouveau alimentés, la cohérence de l'ensemble ne pouvait être assurée notamment par manque de contrôle sur l'alimentation via le réseau électrique du CH. La mise en place d'un second onduleur par salle et disposant des mêmes caractéristiques pourra être à étudier dans le cadre de prochaines évolutions.

3.3.2.2 Les évolutions stratégiques liées à la disponibilité des VM.

Dès la première phase de mise en place, avait été arrêté le choix quant à la disponibilité des données : synchroniser tous les volumes entre les baies. Ce choix n'aura pas été remis en cause.

La principale évolution se situe à ce stade au niveau de la priorisation du redémarrage des machines. Le nombre de machines était au départ réduit et l'ensemble pouvait fonctionner sur une seule machine physique (cas, par exemple, de la disparition totale de la salle SR1). La première approche (voir tableau au tableau X page 90) n'a pas non plus été remise en cause puisqu'il s'agissait des machines définies comme stratégiques pour l'utilisation du dossier patient informatisé (voir synthèse 1.3.2.2 page 26). Toutefois diverses machines sont venues compléter la liste des VM. Donc, toujours dans la continuité de cette approche visant à soutenir les processus métier et notamment de prise en charge du patient tout en tenant compte des contraintes liées à la solution choisie, il a été fait le choix de paramétrer, dans la mesure du possible, un redémarrage automatique pour la machine Sirilog (RIS).

Concernant les autres machines, en lien avec les possibilités d'hébergement des machines physiques mais aussi pour optimiser les performances, le choix a été fait de ne pas les redémarrer systématiquement même si cela pourra être fait manuellement, au cas par cas. L'approche aurait sans doute été différente avec l'offre non retenue lors du marché public.

Le cas de la machine de gestion des commandes de repas (SrvWinrest) peut se débattre. Le redémarrage n'est actuellement pas commandé puisque le choix des machines n'est pas plus fin que ce qui apparaît sur le tableau page suivante et qu'elle ne peut être prioritaire sur les serveurs SrvEnovEAI et Sirilog. Et, malheureusement, aucune optimisation permettant de concentrer plus de machines virtuelles sur un même serveur (pour pallier les cas les plus défavorables) n'est envisageable puisque le provisioning des ressources serveurs ne permet pas de jouer sur les processeurs virtuels alloués.

Tableau XIII : Priorité de redémarrage des VM en phase 2

Nom et description de la VM	Priorité de redémarrage via le dispositif de protection HA
Citrix License Server Virtual Appliance	Restart first
Red Hat Enterprise Linux 5.3 (serveur DPI)	Restart
TSE1 2008 R2 (serveur de présentation)	Restart
AD 2008 x64 (Contrôleur Domaine, DNS, etc.)	Restart
Srvenoveai (EAI pour la gestion des interfaces)	Restart if possible
Sirilog	Restart if possible
PMSIPilot	Do not Restart
SrvQHS	Do not Restart
SrvVirus	Do not Restart
SrvWinrest	Do not Restart
Onduleur RT3000 sur xen1	Do not Restart
Onduleur RT3000 sur xen2	Do not Restart
Onduleur RT3000 sur xen3	Do not Restart

Quant à la machine Sirilog, la non compatibilité avec les XenTools a bien entendu une incidence quant à la surveillance du système par l'hyperviseur mais pas dans la capacité de ce dernier à redémarrer la VM si elle se trouve arrêtée ou si le serveur physique qui la supporte n'est plus disponible.

Finalement, les machines dédiées à la gestion de la protection électrique ne sont pas redémarrées automatiquement par le dispositif de protection HA, d'autant qu'elles sont hébergées et attachées à un serveur bien précis et qu'elles n'ont pas la capacité à être démarrées sur une autre machine. Mais leur présence reste stratégique et leur démarrage est piloté en local par chaque serveur physique dès qu'il est opérationnel.

3.3.3 LES SAUVEGARDES

Les procédures de sauvegarde restaient à ce stade largement à parfaire puisque les solutions mises en place initialement ne permettaient que l'export manuel de VM complète sur un LUN dédié mais toujours sur le même périphérique voire, de manière non structurée sur différentes machines pour le second niveau. Le cycle de sauvegarde de données étaient quant à lui réduit à sa plus simple expression.

3.3.3.1 Planification et stratégie de sauvegarde des données

Jusque-là, les sauvegardes de données reposaient sur deux principes que sont l'export de données sur une machine tierce (7j/7) et la sauvegarde sur bande via des agents installés sur quelques VM. Pour ce dernier cas, nous comptions alors 5 bandes par semaines (du lundi au vendredi) sur deux semaines soit 10 bandes.

C'est ce dernier point qui a fait l'objet d'évolution puisqu'il apparaissait nécessaire de prendre en compte les sauvegardes du week-end mais aussi une problématique d'archivage permettant de retrouver des données plus anciennes. Le cycle a été défini comme suit :

Tableau XIV : Stratégies de sauvegarde de données

Jour	Fonctionnement de la tâche
Lundi Jeu de bandes important puisque la stratégie doit permettre de retrouver les données de chaque lundi pendant deux mois puis de chaque premier lundi du mois pendant 1 an.	Formatage de la bande, sauvegarde complète, vérification des données écrites et éjection en fin de processus.
Mardi, Mercredi, Jeudi (jeu de bandes sur deux semaines paire et impaire)	Formatage de la bande, sauvegarde complète avec exceptions, vérification des données écrites et éjection en fin de processus.

Jour	Fonctionnement de la tâche
Vendredi (jeu de bandes sur deux semaines paire et impaire)	Formatage de la bande, sauvegarde complète avec exceptions et vérification des données écrites. La bande n'est pas éjectée.
Samedi (bandes du vendredi)	Sauvegarde différentielle avec exceptions complétant le support présent et vérification des données écrites. La bande n'est pas éjectée.
Dimanche (bandes du vendredi)	Sauvegarde différentielle avec exceptions complétant le support présent, vérification des données écrites et éjection en fin de processus.

Les machines concernées par ce processus de sauvegarde sont les mêmes que lors de la première phase puisque il n'y avait pas un intérêt direct à sauvegarder selon ce principe les nouvelles VM (VM dédiées à la protection électrique, SrvVirus, SrvWinrest, PMSIPilot, Sirilog) sauf la machine SrvQHS. Suite aux difficultés de virtualisation, la nouvelle installation a été l'occasion de faire reposer la solution logicielle de cette dernière sur un système d'exploitation Windows 2008 R2. L'outil HP Data Protector Express 4 n'étant pas compatible avec celui-ci, les sauvegardes de base de données passent par une phase d'export peut contraignante vu les volumes. Ce point reste toutefois à faire évoluer dès que possible.

3.3.3.2 Sauvegarde des VM complètes

Les sauvegardes de VM complètes ont quant à elles été totalement revues afin de structurer et automatiser les différents processus. D'un point de vue matériel tout d'abord, nous avons complété l'installation par un dispositif reposant sur un NAS connecté au réseau SAN assurant l'externalisation des données sauvegardées sur un autre support que les baies elles-mêmes. Elles représentaient jusque-là une sorte de SPOF dans le

dispositif de protection des données sauvegardées. Ce NAS (en réalité un NUS dont nous n'avons pas exploité les capacités iSCSI) intègre un dispositif de disques redondants (RAID1).

L'opération de sauvegarde précédemment manuelle, ne correspondait pas à l'attente initiale au moins par les aspects suivants :

- L'opération était consommatrice de temps puisqu'il était nécessaire d'en surveiller le déroulement afin de redémarrer au plus vide la VM concernée. De plus la surveillance régulière n'excluait pas un délai de réaction entre la fin de sauvegarde et le redémarrage de la VM.

- Le risque d'erreurs, bien que l'opération n'ait pas été complexe, n'était pas négligeable du fait d'une suite d'opérations discontinue dans un contexte de travail propice aux interruptions pour des opérations de support notamment. Les conséquences pouvaient aussi être importantes puisque les manipulations concernaient des opérations de suppressions de VM ou de paramétrages spécifiques aux VM copiées qui risquaient, en cas d'erreur, de provoquer le démarrage de plusieurs versions d'une même machine en même temps (via la protection par la HA à désactiver par exemple).

- Les opérations n'étaient pas planifiables sans contrainte en dehors des heures de présence des personnels du service informatique puisque non automatisées.

L'opération consistant à exporter les VM vers une machine tierce présentait sensiblement les mêmes travers à l'exception de la surveillance pour, entre autre, la limitation des coupures.

En réponse à cet ensemble de besoins, l'opération a été automatisée en respectant les deux grandes approches suivantes. Tout d'abord, l'automatisation a été faite au travers de scripts se voulant réduits à un ensemble de tâches cohérent et minimal tout en étant assez complet pour laisser l'infrastructure dans un état cohérent en fin d'exécution. Ces scripts ont été réalisés au travers de scripts Linux (l'hyperviseur reposant sur le noyau Suse Linux) intégrant des commandes propres à XenServer. La seconde approche vise à respecter l'autre thème objet de ce mémoire : la haute disponibilité. De celle-ci découle notamment la possibilité d'une panne et l'interchangeabilité des machines physiques. Donc, dans le cadre de la planification des opérations, l'ensemble des serveurs physiques possède des planifications identiques exécutant des scripts localisés dans des emplacements partagés. Pour que les tâches ne s'exécutent pas simultanément de toutes parts, les automatismes intègrent un test permettant d'identifier le rôle du serveur dans le pool afin que seul le pool master lance ces opérations (il sera toujours unique et forcément présent).

Le processus de sauvegarde à froid de l'ensemble d'une VM repose sur deux scripts principaux auxquels il convient de passer en paramètre, lors de l'appel, le nom de la machine virtuelle à considérer. Le premier script, pour peu que l'exécution se fasse sur le pool master, supprime au besoin la précédente sauvegarde et les disques virtuels associés localisés sur les baies SAN dans un LUN dédié, arrête la VM à sauvegarder, lance la nouvelle copie, modifie les paramètres (en particulier de redémarrage et de protection haute disponibilité de la copie une fois faite), redémarre la VM originale sur le serveur le plus disponible. Le second script, toujours réservé à une exécution sur le pool master, procède à la connexion au partage CIFS du NAS, supprime l'éventuel avant-dernier export déjà présent, renomme le dernier en avant-dernier, procède à export de la dernière copie de VM dans le partage avant de le démonter. L'ensemble des opérations

réalisé par ces deux scripts fait l'objet d'un suivi au travers de fichiers d'historique. Ces opérations n'ont pas été intégrées dans un même automatisme afin de pouvoir sauvegarder plusieurs machines de suite dans un délai réduit ; l'export, très long, pouvant quand lui être réalisé à la suite. En fonction des besoins, ces scripts sont appelés individuellement ou au travers d'un troisième se chargeant uniquement de lancer les deux précédents. Pour d'autres besoins, des sauvegardes sous forme d'une liste de sauvegardes suivie d'une liste d'export ont aussi été utilisées.

Finalement, pour tester la qualité des sauvegardes mais aussi plus tard dans le cadre de mises à jour défectueuses, des restaurations ont été faites assurant de la bonne qualité de l'ensemble. Toutefois, sur chaque copie, XenServeur modifie l'adresse MAC de la carte réseau virtuelle. Cela évite les risques induits par la cohabitation accidentelle sur le même réseau de plusieurs machines avec des adresses MAC communes mais pose le problème des restaurations pour certains applicatifs utilisant directement cette information. C'est, par exemple, le cas de la machine PMSIPilot dont la vérification de la clef de licence est basée sur cette donnée. Plutôt que de l'extraire de la machine source pour la paramétrer sur la copie, j'ai fait le choix de tenir à jour une table de ces adresses pour chacune des machines virtuelles comme nous le faisons déjà pour l'ensemble des machines physiques (gestion des VLAN et démarrages à distance). Là encore, s'il fallait le rappeler, les tests en réel restent nécessaires. Il y en a donc, bien entendu eu d'autres dont, par exemple, la restauration d'un export de la machine DIS sur l'infrastructure Citrix de l'éditeur en charge de sa maintenance.

Là encore, il persiste des points d'amélioration que nous verrons à la suite.

3.4 POINTS D'AMELIORATION ET PERSPECTIVES

Ont précédemment été listées, à l'occasion de chacun des thèmes, les limites auxquelles nous nous sommes arrêtés pour chacun d'eux. D'une manière plus globale, il apparaît nécessaire de poursuivre le développement de cette infrastructure en continuant de tendre vers les objectifs ayant conduit à la mise en place de cette solution et des solutions connexes.

Dans ce cadre, l'élimination des « Single Point Of Failure » et des risques liés à l'environnement de fonctionnement parait être prioritaire. En premier lieu, il faudra se concentrer sur la sécurisation de la machine de management qui héberge la FOM nécessaire au quorum pour l'élection de baies SAN, l'espace disque partagé pour les scripts et particulièrement ceux nécessaires au dispositif de protection électrique. Cette sécurisation passe par la fiabilisation de la machine et son déport de la baie SR1 où les dysfonctionnements engendrés par l'environnement sont, en l'état actuel, commun avec une moitié de l'infrastructure. L'évolution du NAS devrait aussi être calquée sur cette approche.

Dans le même esprit, le stockage devra être revu sur plusieurs aspects. Tout d'abord prioriser la reconstruction de données sur l'utilisation par les VM afin de retrouver au plus vite un niveau normal de tolérance aux pannes. Par ailleurs, le découpage des LUN pour permettre les sauvegardes de VM par la technique du « Fast clone » moins contraignante pour les utilisateurs et l'utilisation des fonctions de Thin Provisioning devront aussi être programmées.

L'exécution de VM avec systèmes d'exploitation non compatibles avec l'hyperviseur est aussi source de questionnement quant à la pérennité de ces dernières et nécessite une gestion spécifique. En effet, lors des sauvegardes de VM ou d'arrêt piloté par les

onduleurs, l'hyperviseur n'est pas en mesure d'arrêter proprement ces machines. Se pose également le problème de déplacement à chaud voire de performance : ces VM ne bénéficiant pas de la paravirtualisation. La stratégie retenue sera calquée sur la stratégie métier qui conduira à maintenir ou non ces applications. Si tel est le cas, leur inévitable évolution conduira à la remise en cause des machines virtuelles associées.

Finalement, les outils de sauvegarde capables de prendre en charge les systèmes d'exploitation récents voire l'automatisation du redémarrage de l'infrastructure après une défaillance électrique seraient aussi à mettre en place.

CONCLUSION

La virtualisation de serveurs n'est pas un concept nouveau mais s'est développée au point de devenir aujourd'hui une composante essentielle de la plupart des stratégies. Associée au concept de haute disponibilité, elle permet de mutualiser celui-ci à l'ensemble des serveurs de l'entreprise. Serveurs dont le nombre croît au rythme de l'informatisation des processus de l'organisation et dont la disponibilité et l'agilité deviennent primordiales. C'est à mon sens cet aspect qui fait aujourd'hui le succès de la virtualisation en haute disponibilité dans les PME et peut-être également le succès, demain, de l'externalisation de tout ou partie du système d'information de l'entreprise, ou de la virtualisation du poste de travail.

Développées au sein du Centre Hospitalier du Blanc, ces approches et techniques nous ont permis, à ce jour, de répondre aux besoins de disponibilité exigés par le processus médical et paramédical de prise en charge du patient. En effet, bien qu'une comptabilisation exacte des périodes d'indisponibilités imprévues n'ait pas été tenue depuis la mise en production, il est toutefois possible d'avancer que le dispositif a joué son rôle pour l'ensemble des environnements hétérogènes consolidés sur l'hyperviseur et identifiés comme nécessaires.

Toutefois, de par les éléments mis de côté pour des raisons économiques mais aussi dans le cadre d'une recherche perpétuelle d'amélioration de la qualité des installations, un certains nombre de points devront, encore à ce jour, faire l'objet d'amélioration. Améliorations d'ordre technique ou organisationnel pour améliorer directement le dispositif ici présenté, à commencer par la recherche des SPOF et l'étude de la reprise après sinistre, mais aussi amélioration quant à la conduite de projet qui doit se reposer sur l'utilisation de bonnes pratiques reconnues et sur la démarche d'amélioration continue qui

caractérise les processus qualité. Finalement, le dernier thème d'amélioration, sans qu'il ne s'agisse d'une présentation chronologique devra être, en premier lieu, la mise en place d'un Plan de Reprise d'Activité et, de par la mise en commun des compétences, l'étude des processus de prise en charge du patient et la rédaction d'un Plan de Continuité d'Activité. Il faut rappeler que la plateforme n'est pas infaillible et qu'elle ne traite qu'un type de causes d'indisponibilité.

Au-delà, le projet ici présenté s'intègre dans un environnement qui se doit de répondre aux mêmes exigences dès qu'il participe à la même mission et dans un périmètre équivalent. Pour cette raison, j'ai recherché à conjuguer au mieux ces différents projets dans le cadre d'une approche globale et systémique. Ce projet, et plus largement l'ensemble des projets ici évoqués, nous montrent qu'il était indispensable de prendre le recul nécessaire pour appréhender l'ensemble de la problématique, prioriser les choix et ne pas se focaliser sur des aspects seulement informatiques sans prendre en compte l'ensemble de l'organisation.

Le thème évoqué au travers de ce document n'est pas seulement un projet bien délimité dans le temps, mais plus une étape importante d'un processus continu d'optimisation des ressources et d'amélioration de la qualité du service rendu. Il apparaît vraisemblable que cette approche, perfectible dans sa mise en application, sera celle que l'on me demandera de poursuivre à l'occasion des nouvelles missions qui me seront confiées dans le cadre du rapprochement des structures hospitalières opéré à l'occasion de la mise en place de Communauté Hospitalière de Territoire de l'Indre.

BIBLIOGRAPHIE

[1] CODE DE LA SECURITE SOCIALE, *Décret n°2005-1023, 2009,* http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000631121

[2] Code De La Santé Publique, *Décret n° 2003-462 du 21 mai 2003 relatif aux dispositions réglementaires des parties I, II et III du code de la santé publique, Article R.1112-2,* 2003, http://legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000412528.

[3] Centre Hospitalier du Blanc, *Livret d'accueil des stagiaires et collaborateurs occasionnels,* 2008.

[4] AGENCE TECHNIQUE DE L'INFORMATION SUR L'HOSPITALISATION, *PMSI : analyse de l'activité et des aires de recrutement des établissements de santé*, 2012, http://carto.atih.sante.fr/cp.php?reg=24.

[5] Code De La Santé Publique, *Article R6141-16*, 2005, http://www.legifrance.gouv.fr/affichCodeArticle.do?idArticle=LEGIARTI000006917396.

[6] Centre Hospitalier du Blanc, *Collection de documents et photographies à disponibilité des personnels d'encadrement*, 2011.

[7] Centre Hospitalier du Blanc, *Rapport d'activité et de gestion 2010*, 2011.

[8] OBSERVATOIRE DES SYSTEMES D'INFORMATION DE SANTE, *Budgets SI : Synthèse comparative 2009*, 2012, https://o6.sante.gouv.fr.

[9] Laboratoire d'informatique médicale de la Faculté de Médecine de l'Université de Rennes 1, *Missions du D.I.M.,* 2011, http://www.med.univ-rennes1.fr/plaq/dim/missions.html

[10] Téléphonie Française du Centre, *Documentation technique - Marché Interconnexion sites du Centre Hospitalier de Le Blanc*, 2007.

[11] ALCATEL-LUCENT ET CH DU BLANC, *Alcatel Omnivista 2500 version 3.4.2GA*, 2011, Menu « Topology », « Logical Network ».

[12] INTERNET ENGINEERING TASK FORCE, *RFC1918 : Address Allocation for Private Internets*, 1996, http://tools.ietf.org/html/rfc1918

[13] ENOVACOM, *Urbaniser votre système d'échanges électroniques pour répondre à vos nouveaux enjeux*, 2012, http://www.enovacom.fr/fr/vos-projets/echanges-electroniques-professionnels-edi

[14] RENE J CHEVANCE, *Systèmes à haute disponibilité : Concepts*, Ed. Techniques de l'Ingénieur, 1999

[15] AMD, *Les clés de la construction d'une solution de virtualisation robuste*, 2012, https://rapidrequest.emediainternational.fr/2.aspx?523191.UEOEYXBG.4922

[16] CNAM, *Unité d'Enseignement ENG110*, 2011, Paris.

[17] TANEJA GROUP, *Une solution de virtualisation VMware simple, économique et éprouvée pour protéger les PME*, 2011, http://download3.vmware.com/elq/MZ/SMB_phase3/assets/fr/BCDR-Taneja-Group-FR_March_2011.pdf

[18] VMWARE, *Établir le dossier commercial de la virtualisation*, 2011, http://www.lemondeinformatique.fr/livre-blanc/etablir-le-dossier-commercial-de-la-virtualisation-1188.html

[19] SEARCHDATACENTER.COM, *L'Etat de la virtualisation et du cloud computing en 2011*, 2011, http://www.lessourcesit.fr/letat-de-la-virtualisation-et-du-cloud-computing-en-2011/

[20] CALIPIA CONSULTING, *Cahier des charges générique pour un projet de virtualisation de serveurs*, 2011, http://fr.slideshare.net/ZDNet_France/pme-dmarrez-votre-projet-de-virtualisation

[21] CALIPIA CONSULTING, *10 conseils pour réussir votre projet de virtualisation de serveurs*, 2011, http://fr.slideshare.net/projectsi/conseils-virtualisation-serveurs

[22] BESLUAU EMMANUEL, *Management de la continuité d'activité,* Ed.Eyrolles, 2010.

[23] VISION SOLUTIONS, *Le guide essentiel de la reprise après sinistre. Comment assurer la continuité informatique et fonctionnelle*, 2010, http://www.visionsolutions.com/world/francaiseu/WPD-EssentialGuideDR-FR.aspx?%20CampaignId=701600000005Dc1&WhitePaper=WPD_EssentialDR_FR.pdf

[24] RENE J CHEVANCE, *Systèmes à haute disponibilité : Solutions*, Ed. Techniques de l'Ingénieur, 1999

[25] JACQUES PEPING, *Architecture de systèmes de stockage*, Ed. Techniques de l'Ingénieur, 1997

[26] JACQUES PEPING, *Stockage de données en réseau SAN*, Ed. Techniques de l'Ingénieur, 2003

[27] Wikipedia.org, *RAID*, 2013, http://fr.wikipedia.org/wiki/RAID_(informatique)

[28] KENNETH HESS, AMY NEWMAN, Virtualisation en pratique, 2010, ed. Pearson Education France

[29] BERTIL FOLLIOT, GAËL THOMAS, Virtualisation logicielle : De la machine réelle à la machine virtuelle abstraite, Ed. Techniques de l'Ingénieur, 2009

[30] Centre Hospitalier du Blanc, *MARCHE N° 10-01-01 / infrastructure de virtualisation / Cahier des Clauses Techniques Particulières*, 2011.

[31] Hewlett-Packard Development Company, *Building High Performance High Availability IP Storage Networks with SANiQ*, 2009, http://h10032.www1.hp.com/ctg/Manual/c01750150.pdf

TABLE DES ILLUSTRATIONS

LISTE DES TABLEAUX

Sécurisation de l'environnement matériel d'exécution d'un dossier patient informatisé. Mémoire d'Ingénieur CNAM, 2013.

RESUME

L'informatisation toujours croissance du processus de prise en charge médicale et paramédicale du patient, nécessite une disponibilité maximum des ressources informatiques, pour assurer la continuité inhérente à ce type d'activité.

L'hôpital du Blanc souhaite se lancer dans la démarche d'informatisation du dossier patient, à commencer par le circuit du médicament. En parallèle du choix de la solution logicielle, sont menés, au travers d'une approche globale, différents projets de mise à niveau de l'infrastructure afin d'en permettre l'exploitation par les utilisateurs. Le présent mémoire se focalise sur la virtualisation des serveurs et du stockage en haute disponibilité.

Après avoir présenté le contexte du projet puis les concepts et solutions techniques permettant d'atteindre les objectifs, ce mémoire présente, au cours d'un troisième chapitre, les différentes étapes de la mise en place.

Mots clefs : Virtualisation, Haute disponibilité, Serveur, Stockage, P2V, Hyperviseur, SAN, SPOF.

SUMMARY

The ever growing computerization of the process of medical and paramedical care of the patient requires maximum availability of IT resources, to ensure the continuity inherent in this type of activity.

The hospital of Le Blanc wants to start the process of computerization of patient records, starting with the medication. Alongside the choice of the software solution are conducted, through a holistic approach, different projects of upgrading infrastructure to allow operation by users. This paper focuses on High Availability server and storage virtualization.

After presenting the context of the project and the technical concepts and solutions to achieve the objectives, this paper presents, in a third chapter, the various stages of implementation.

Keywords : Virtualization, High Availability, Server, Storage, P2V, Hypervisor, SAN, SPOF.