



Multimedia infrastructure optimisation backup : migration de l'outil de sauvegarde, refonte des mécanismes de sauvegarde et virtualisation

Florent Matzinger

► To cite this version:

Florent Matzinger. Multimedia infrastructure optimisation backup : migration de l'outil de sauvegarde, refonte des mécanismes de sauvegarde et virtualisation. Cryptographie et sécurité [cs.CR]. 2011. dumas-01224103

HAL Id: dumas-01224103

<https://dumas.ccsd.cnrs.fr/dumas-01224103>

Submitted on 4 Nov 2015

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

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

CONSERVATOIRE NATIONAL DES ARTS ET METIERS

CENTRE REGIONAL ASSOCIE DE STRASBOURG

MEMOIRE

Présenté en vue d'obtenir

Le Diplôme D'INGENIEUR C.N.A.M.

En

INFORMATIQUE

Par

Florent MATZINGER

MULTIMEDIA INFRASTRUCTURE OPTIMISATION BACKUP

**MIGRATION DE L'OUTIL DE SAUVEGARDE, REFONTE DES MECANISMES DE SAUVEGARDE ET
VIRTUALISATION**

Soutenu le 28/01/2011

JURY

Président : I. WATTIAU

Membres : J. RABITA

F. ALBERT

R. REIBEL

C. KLEINPETER

Résumé

Ce document, intitulé MULTIMEDIA INFRASTRUCTURE OPTIMISATION BACKUP, décrit l'optimisation des sauvegardes de la plate-forme de service Mobile Data Service Platform chez l'opérateur de télécommunications France Telecom.

Cette optimisation s'est effectuée à travers 3 axes :

- la migration de l'outil de sauvegarde,
- la refonte des mécanismes de sauvegardes de bases de données,
- la virtualisation des sauvegardes.

Ce projet d'optimisation et de consolidation d'infrastructure a été mené de septembre 2007 à mars 2009 grâce à la méthodologie de conduite de projet Time To Market et a permis d'obtenir les résultats suivants :

- migration de près de 1000 serveurs,
- mutualisation et consolidation de l'infrastructure de sauvegarde NBU et des bibliothèques virtuelles,
- libération de l'infrastructure de sauvegarde TiNa et des bibliothèques physiques,
- simplification des mécanismes de sauvegardes de bases de données.

Mots-clés

Migration, NetBackup, sauvegarde, TiNa, virtualisation, VTL.

Sum-up

This document, named MULTIMEDIA INFRASTRUCTURE OPTIMISATION BACKUP, describes the backup optimization of the Mobile Data Service Platform from the telecommunications operator France Telecom.

This optimization operated through 3 axes:

- backup software migration,
- database backup procedures consolidation,
- backup virtualization.

This optimization and consolidation project was led from September 2007 to March 2009 under the Time To Market project leading methodology and allowed the following results:

- migration of about 1000 servers,
- NBU backup infrastructure and virtual libraries mutualization and consolidation,
- TiNa backup infrastructure and physical libraries decommissioning,
- database backup procedures simplification.

Keywords

Backup, migration, NetBackup, TiNa, Virtualization, VTL.

Sommaire

1. L'environnement – l'entreprise d'accueil	6
1.1. Historique du groupe France Telecom	6
1.2. Évolution du groupe France Telecom : la marque Orange	6
1.3. Chiffres clés du groupe France Telecom	7
1.4. Organigramme	8
1.5. La Direction Direction des Plates-formes de Service (DPS)	9
1.6. L'entité Infrastructures Techniques d'Entreprise (ITE)	10
1.7. La Direction de projet Architecture de Stockage et Sauvegarde (AS)	10
1.8. Le pôle sauvegarde - l'équipe MOE sauvegarde	11
1.9. Le candidat	12
2. Le projet et son contexte	13
2.1. Les services mobiles	13
2.2. La plate-forme Mobile Data Service Platform (MDSP)	14
2.3. Le programme Multimedia Infrastructure Optimisation (MIO)	16
2.4. Le projet MIO Backup	17
a) Objectifs	17
b) Enjeux de la solution de sauvegarde déployée	18
c) Contexte de la sauvegarde actuelle	18
d) Cible	18
e) Qualité de service	19
2.5. Périmètre initial	19
3. L'état de l'art	20
3.1. Le but de la sauvegarde	20
3.2. Les principes de la sauvegarde	20
3.3. L'évolution de la sauvegarde	27
3.4. La virtualisation	28
3.5. La déduplication	30
a) Le principe	30
b) L'algorithme	31
c) Évolution vers le stockage primaire	32
d) Adoptée par toute l'industrie du stockage	32
3.6. Conclusion	33
4. Organisation du projet	34
4.1. Méthodologie de la conduite du projet	34
a) les objectifs	34
b) les facteurs clés de succès de TTM	35
c) les phases et jalons	36
4.2. Les passages obligatoires du projet	37
a) Le CVAT	37
b) Le CI2A	38
c) Le COGIT	38

4.3.	Organisation hiérarchique	38
a)	Services en charge de la gestion du projet	39
b)	Services en charge de l'exploitation	40
4.4.	Activités confiées	41
5.	Analyse de l'existant	42
5.1.	Audit du périmètre à sauvegarder	43
5.2.	L'environnement technique	44
5.3.	Le découpage en environnement et DMZ	45
5.4.	L'intégration dans l'environnement cible : les IAS	46
5.5.	Audit du SAN de backup	47
5.6.	Contraintes et risques du projet	49
6.	Définition de la solution technique	52
6.1.	Architecture générale	52
6.2.	Calculs capacitifs	55
6.3.	Choix des éléments de l'infrastructure de sauvegarde	60
6.4.	Intégration dans l'IAS	67
a)	réseau IP	69
b)	réseau SAN	70
	Scénario b1 : Interconnexion des SAN de backups Orange et IAS	70
	Scénario b2 : Double attachement des Storage Nodes Tina	72
	Scénario b3 : Double attachement des DXi	73
6.5.	Livrables	75
6.6.	Rédaction du DAT et passage en PCVAT/CVAT	75
6.7.	Sauvegarde des bases de données	79
a)	Oracle	79
b)	DB2	83
c)	MySQL	83
6.8.	Reporting - Supervision des sauvegardes	84
6.9.	Découpage du projet en phases - mise au point des procédures générales	85
	Phase 0 : État des lieux	85
	Phase 1 : Déploiement	86
	Phase 2 : Migration	87
	Phase 3 : Fin du projet	88
7.	Mise en œuvre de la solution	89
7.1.	Planning	89
	Phase 1 : Déploiement	90
	Phase 2 : Migration	90
7.2.	Déploiement de l'infrastructure de sauvegarde	91
7.3.	Pilote	97
7.4.	Principe d'installation des packages clients et migration des clients	98
a)	Pré-requis	98
b)	Installation	98

7.5.	Principe d'installation des packages SAN media serveurs et migration des storage nodes.....	99
a)	Pré-requis	99
b)	Installation	99
c)	Activation - post installation.....	99
7.6.	Principe d'installation des agents de base de données et migration des sauvegardes des bases de données	100
7.7.	Problèmes rencontrés	100
a)	Changements d'équipe.....	100
b)	Xinetd	101
c)	Retard dans la rédaction des EB	101
d)	Interactions avec d'autres projets	101
e)	Problème de restauration des bases de données Oracle en cluster.....	102
8.	Fin du projet.....	103
8.1.	Bilan de projet.....	103
a)	Architecture mise en place.....	103
b)	Synthèse globale	104
c)	Gestion des risques	104
d)	Bilan pilotage projet	105
8.2.	Gains	105
8.3.	Mise en place de la gestion récurrente	107
8.4.	Transfert de compétence aux équipes France Telecom	108
8.5.	Qualité de service - Mise en place de la BQN	108
9.	Conclusion	111
9.1.	Bilan personnel - améliorations	111
9.2.	Apport du cursus CNAM	111
9.3.	Remerciements.....	112
10.	Bibliographie.....	113
11.	Glossaire	114

1. L'ENVIRONNEMENT – L'ENTREPRISE D'ACCUEIL

Je vais exposer dans ce premier chapitre le groupe France Telecom (FT), la Direction de la construction des Plates-formes de Services et des commandes réseaux (DPS), l'entité Infrastructures Techniques d'Entreprise (ITE), la Direction de projet Architecture de Stockage (AS) et le pôle sauvegarde auquel je suis rattaché en tant que prestataire externe.

1.1. Historique du groupe France Telecom

L'histoire de France Telecom est étroitement liée au réseau de télécommunications national. C'est pourquoi il faut remonter à l'apparition du télégraphe de Chappe pour expliquer l'évolution de cette entreprise de sa création jusqu'à aujourd'hui.

En 1837, un monopole est décrété sur le réseau de télégraphie optique (télégraphe de Chappe), puis son administration est regroupée avec les services postaux au sein du ministère des Postes et Télégraphes (PT) en 1878.

Avec la seconde guerre mondiale, l'État saisit l'importance des télécommunications et crée la direction des télécommunications au sein des Postes, Télégraphes et Téléphones (P&TT) en 1941. Un centre national d'études des télécommunications (CNET) fut même créé en 1944. À partir des années 70, une forte demande des équipements téléphoniques obligea la Direction Générale des Télécommunications (DGT) à avoir un fonctionnement plus indépendant de celui de la Poste.

La DGT prend le nom de France Telecom en 1988 et devient exploitant autonome de droit public le 1er Janvier 1991. À partir de la fin des années 70 et du début des années 80, le réseau national des télécoms se voit suffisamment développé géographiquement et techniquement (câble, satellite) afin que la DGT, avec sa maîtrise des transmissions numériques, puisse ouvrir Transpac en 1978 (1er réseau de transmission de données par paquets), puis propose le Minitel en 1983 et enfin lance le premier satellite de télécommunications français en 1984. Le service Radiocom 2000, qui permet en 1986 de communiquer avec un correspondant en voiture, annonce le début de la radio télécommunication. En 1991, l'ouverture du réseau GSM va permettre le développement du téléphone mobile numérique européen Orange.

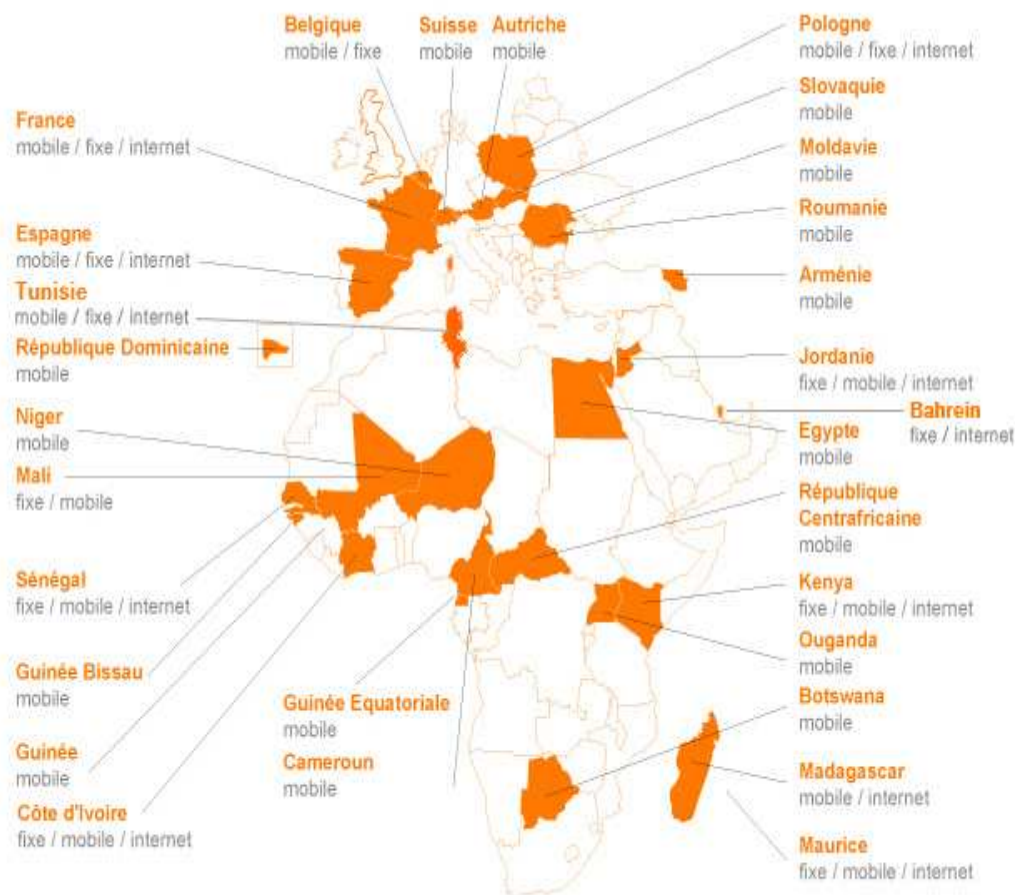
1.2. Évolution du groupe France Telecom : la marque Orange

Depuis, France Telecom décide de se positionner comme un opérateur de télécommunications globales et s'enrichit de services et de produits Internet (Wanadoo), de réseaux professionnels (Equant), de multimédia numérique comme la télévision numérique (MaLigne TV). Avec ce développement des services, France Telecom choisit de simplifier son fonctionnement pour ses clients. Ainsi, suite au rachat de Orange UK en août 2000, la marque commerciale «France Telecom» change de nom et devient **Orange** le 1er Juin 2006. Orange représente désormais la marque unique du Groupe pour l'Internet, la télévision et le mobile en France, au Royaume-Uni, aux Pays-Bas et en Espagne, et Orange Business Services la marque des solutions et services commercialisés auprès des entreprises et des administrations dans le monde. Ce changement de marque a aussi permis le regroupement des services clients, portails Internet et boutiques physiques

sous le nom d'Orange. Seules les offres de téléphonie classique conservent la marque France Telecom en France.

Aujourd'hui France Telecom ne se positionne plus uniquement en fournisseur de réseau mais également en tant qu'**opérateur international de services intégrés**, c'est-à-dire en fournissant aussi bien les services que les réseaux qui permettent d'y accéder dans le monde entier. Certains de ces services sont hébergés sur la plate-forme **MDSP** sur laquelle porte mon projet et que je présenterai plus en détail au chapitre 2.2.

Cette carte représente les différentes activités d'Orange que sont le fixe, le mobile et l'internet dans le monde :



Nous offrons des services pour le grand public dans 32 pays et pour les entreprises dans 166 pays.

Illustration 1 : la marque Orange dans le monde

1.3. Chiffres clés du groupe France Telecom

Le Groupe France Telecom est aujourd'hui l'un des principaux opérateurs de télécommunications au monde. En 2007 (année de lancement du programme MIO sur lequel porte mon mémoire), le Groupe France Telecom a réalisé un chiffre d'affaires consolidé de 53 milliards d'euros et sert près de 170 millions de clients sur les cinq continents.

Sa couverture mondiale est assurée par :

- le plus grand Réseau voix/données Sans Couture (RSC📖) au monde (couvre 220 pays),
- des réseaux de téléphonie mobile dans 32 pays,
- une assistance disponible dans 166 pays.

La marque Orange se classe parmi les toutes premières marques mondiales :

- un des leaders mondiaux sur le marché des mobiles, n°3 en Europe, n°1 en France,
- 3ème fournisseur de lignes fixes dans le monde, n°1 en Europe,
- 4ème fournisseur d'accès ADSL dans le monde, n°1 en Europe.

Orange obtient régulièrement de nombreux prix qui assoient sa position de leader :

- reconnaissance du marché :
 - o meilleur opérateur mondial World Communication Awards 2006, 2007, 2008 et 2009,
 - o meilleur opérateur mobile World Communication Awards 2006, 2008 et 2009,
 - o meilleur initiateur de changement World Communication Awards 2008 et 2009.
- satisfaction du client :
 - o Platinum Award du Meilleur Opérateur Mondial Telemark 2010.
- innovation :
 - o prix «best innovator» A.T Kearney / Les Echos 2007.
- développement durable :
 - o leader du secteur des télécommunications dans le «Green Quadrant» Verdantix 2009.

et des certifications ISO :

- o ISO 15408 sécurité du réseau international IP VPN 2008,
- o ISO 20000 gestion des services 2009,
- o ISO 9001 gestion de la qualité 2009,
- o ISO 27001 système de gestion de la sécurité 2009.

1.4. Organigramme

Ci-dessous une présentation de l'organigramme du groupe France Telecom en 2007 (année de lancement du programme MIO sur lequel porte mon mémoire), concentré sur la direction Division Réseaux, Opérateurs & Système d'Information (**RO&SI**) et ses principales entités avec lesquelles j'ai travaillé dans le cadre de ce projet, puis un zoom jusqu'à l'équipe **DPS/ITE/AS/SAUEGARDE** à laquelle je suis toujours rattaché aujourd'hui :

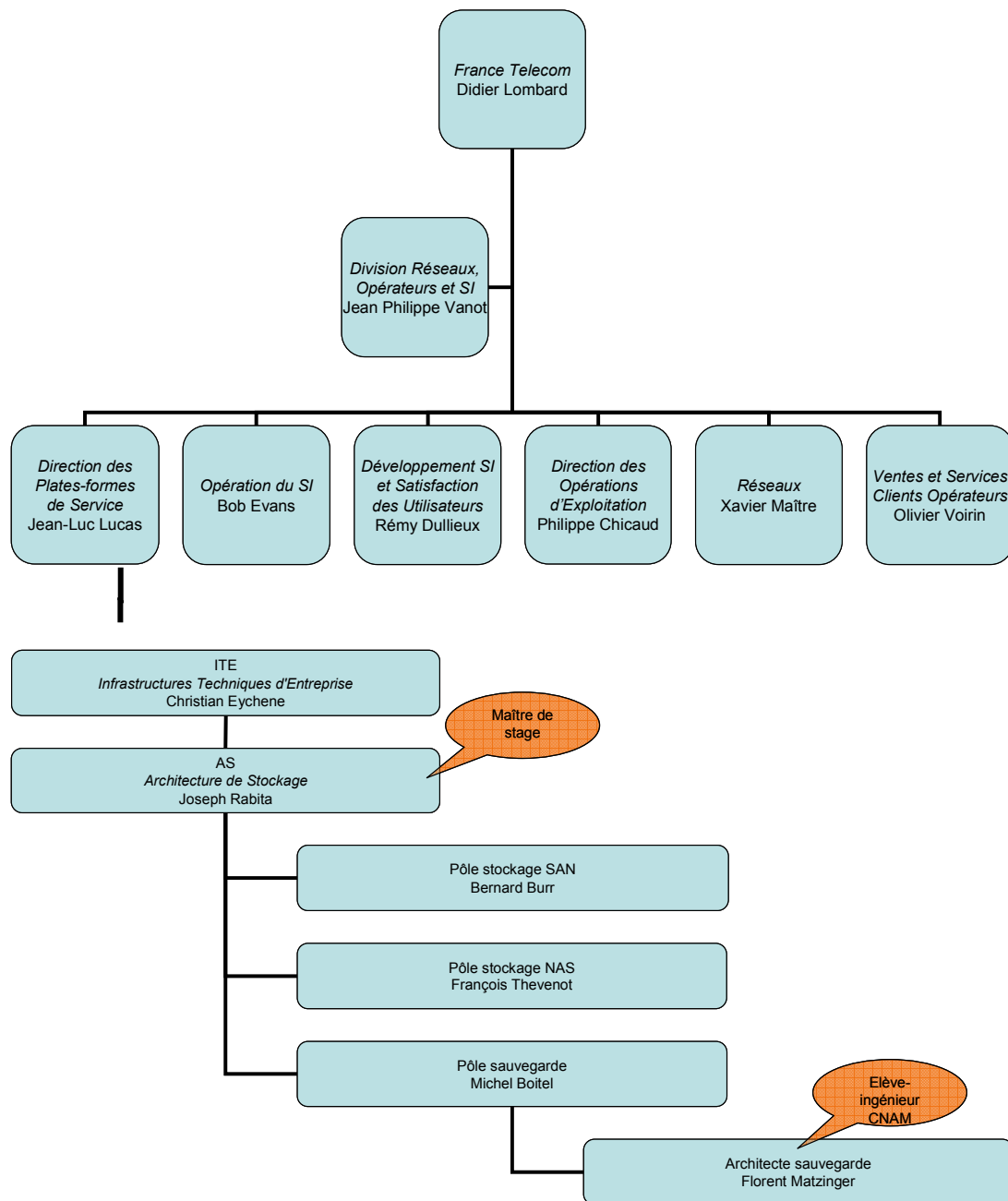


Illustration 2 : mon positionnement dans l'organigramme FT

1.5. La Direction Direction des Plates-formes de Service (DPS)

La Direction de la construction des Plates-formes de Services et des commandes réseaux (**DPS**) au sein de la Division Réseaux, Opérateurs & Système d'Information (**RO&SI**) comprend environ 680 experts sur une dizaine de sites en France.

Les missions de DPS sont :

- **Piloter le développement et la mise en œuvre, pour l'ensemble du groupe, des plates-formes industrielles** fournissant aux clients finaux (résidentiels, mobiles, entreprises et groupe France Telecom) les services Réseaux et SI de l'opérateur intégré France Telecom (exemples : services sur ADSL, messagerie, VoIP, visiophonie, portails, carnets d'adresses, services multimédia tous supports...),

- **Concevoir, développer et maintenir les composants techniques pour le système d'information du groupe France Telecom** nécessaires à la mise en œuvre d'une infrastructure solide et normalisée,
- **Assurer l'intégration** et la métrologie des applications métiers du système d'information du groupe France Telecom afin d'améliorer leur qualité de service et les optimiser.

1.6. L'entité Infrastructures Techniques d'Entreprise (ITE)

L'entité Infrastructures Techniques d'Entreprise a pour mission de réduire les coûts par consolidation et virtualisation des applications et des Datacenter. Elle doit maintenir la qualité de service sur l'ensemble des infrastructures du Système d'Information tout en réduisant les coûts de maintenance. Elle doit développer l'infrastructure IT pour le groupe. Elle assure la construction pour ses clients entreprises des infrastructures de communication vendues par sa filiale Orange Business Services (OBS). **Elle se charge de fournir aux plates-formes de services et à l'IT les infrastructures** de production, les outils de supervision, d'administration et de qualité de service. Elle s'occupe d'accroître la sécurité du poste de travail et de développer la supervision de la sécurité sur les plates-formes de service notamment.

Afin de pouvoir concevoir, développer et maintenir les composants techniques pour les plates-formes de service et le système d'information du groupe France Telecom nécessaires à la mise en œuvre d'une infrastructure solide et normalisée, l'entité ITE est organisée selon le schéma suivant :

- **Pôle AS : Architecture Stockage et Sauvegarde**
- Pôle COSD : Corporate Office Systems Delivery
- Pôle ECV : Expertise, Consolidation & Virtualisation
- Pôle ISRS : Infrastructures et Services Réseau Sécurité
- Pôle MSSOI : Mass Storage & Service Oriented Infrastructure
- Pôle OES : Outil d'exploitation & Sécurité
- Pôle OSIS : Office System Infrastructure & Service
- Pôle PAR : Platon, Activité Récurentes
- Pôle SAQS : Supervision, Administration & Qualité de Service

1.7. La Direction de projet Architecture de Stockage et Sauvegarde (AS)

La Direction de Projet Architecture de Stockage et **Sauvegarde** a la responsabilité de la maîtrise d'œuvre des domaines «stockage» et «sauvegarde». Elle est divisée en trois services ou pôles, à savoir le pôle Stockage SAN (Storage Area Network), le pôle NAS (Network Attached Storage) et le pôle Sauvegarde. La Direction de Projet AS est force de proposition sur l'urbanisme, garante de la cohérence de ce domaine, et a pour mission **la maîtrise d'œuvre sur les solutions de sauvegarde**, de SAN et de NAS.

À ce titre elle assure :

- l'ingénierie sur les technologies de stockage SAN et NAS,
- **la qualification et l'ingénierie sur les technologies de sauvegarde,**
- **les préconisations et prescriptions pour le groupe France Telecom,**
- l'expertise de dernier niveau,
- **l'accompagnement des projets SI et PFS,**
- l'accompagnement des réponses aux appels d'offre d'Orange Business Services (OBS).

1.8. Le pôle sauvegarde - l'équipe MOE sauvegarde

Le pôle sauvegarde, au sein duquel j'exerce, est en charge de la protection de la donnée. Une présentation de l'art de la sauvegarde sera détaillée au chapitre 3 de ce mémoire.

Nos activités au sein de DPS/ITE/AS sont les suivantes :

- Ingénierie infrastructure virtualisée du programme Ecocenter (VCB, autre solution, PRA, spécifications...) *détaillée au paragraphe suivant*
- Ingénierie VTL et robotiques physiques
- Ingénierie des outils de sauvegarde (TiNa et NBU)
- Ingénierie PRA infrastructure de sauvegarde
- Rédaction des prescriptions de sauvegarde (centre de compétences)
- **Ingénierie et déploiement de solution pour des projets spécifiques** (y compris les clients externes)
- Veille technologique permanente sur l'infrastructure (VTL, déduplication à la source ou à la cible ...)
- Étude et proposition de solutions alternatives à NBU et VCB
- MOE sur les outils connexes facteurs d'amélioration de la qualité de service et facilitant la pro- activité
- Accompagnement des projets PFS (suivi et validation des licences et des plans de sauvegarde)
- Support de dernier niveau si non réponse du fournisseur (hors chaîne de soutien)
- Support de l'infrastructure de Sauvegarde
- Skill Center domaine Sauvegarde
- Gestionnaire produit Platon (garant des qualifications NBU, du référentiel documentaire commun)
- **Suivi et proposition d'évolution sur les infrastructures de sauvegarde**
- Responsabilité de la Qualité de Service des solutions d'infrastructure de sauvegarde
- **MOE Sauvegarde en IAS/MDSP «RUN»**

Ci-dessous quelques chiffres pour donner un ordre d'idée du périmètre :

- à la DOSI (le SI interne) : 5000 serveurs sauvegardés, 3Po de données sauvegardées par mois,
- dans les IAS (les plate-formes de service) : 2000 serveurs sauvegardés, 1Po de données sauvegardées par mois.

Le but du programme Ecocenter est de construire une nouvelle architecture de production pour les Datacenter en s'appuyant sur les objectifs suivants :

- Bâtir une infrastructure standard allouant dynamiquement les ressources et des services aux applications aux services hébergées en fonction de leurs besoins et de leur SLA.
- Rompre le lien statique entre une application ou un service et des ressources physiques spécifiques.
- Consolider les parcs de services applicatifs et techniques sur cette infrastructure afin de réduire le nombre de ressources physiques et les coûts liés à leur exploitation.
- Transformer les processus – du développement à la production – pour améliorer le TTM (*voir chapitre 4.1*) et réduire les coûts, en s'appuyant sur cette nouvelle infrastructure.
- Transformer les processus d'exploitation de manière à bénéficier de l'apport des nouvelles technologies mises en œuvre.

Ci-dessous un tableau présentant nos activités récurrentes :

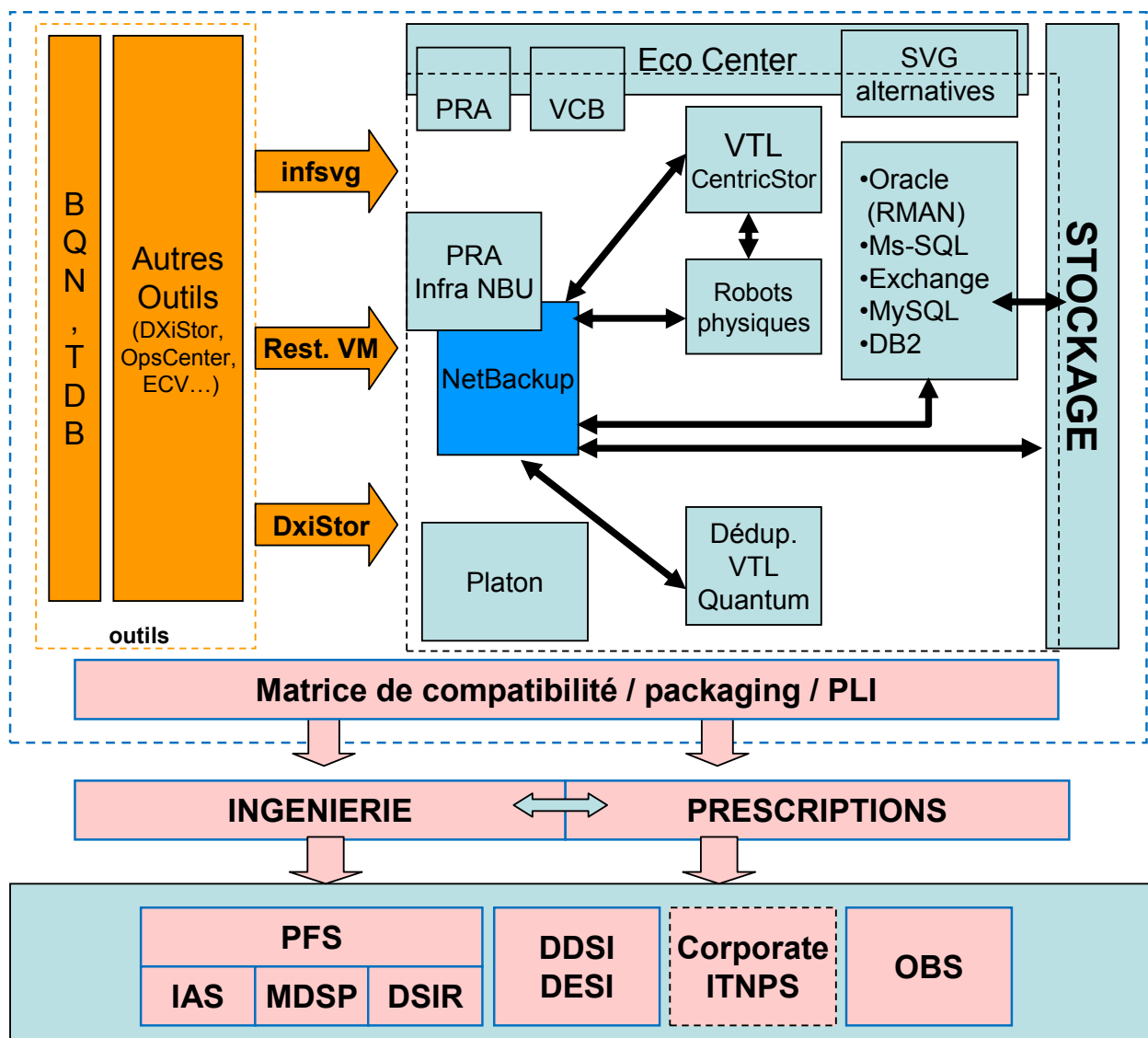


Tableau I : matrice des activités de la MOE sauvegarde

1.9. Le candidat

J'occupe depuis novembre 2007 une fonction d'**expert sauvegarde** au sein de l'équipe DPS/ITE/AS, dont mon maître de stage, M. Joseph Rabita, est le **directeur de projet**. J'assure les rôles de **MOE**, **architecte** et expert support niveau 3, voire chef de projet adjoint dans le cadre du projet confié.



Il est important de noter que **j'ai participé à un projet similaire** au sein du groupe France Telecom en qualité d'architecte et expert support niveau 3 pour la migration TiNa vers NBU et virtualisation sur CentricStor auprès de la Direction des Opérations du Système d'Informations (DOSI). Cette expérience a été déterminante pour ma candidature et m'a été très utile pour mener ce projet à bien.



2. LE PROJET ET SON CONTEXTE

2.1. Les services mobiles

Comme présenté au chapitre 1.2, France Telecom ne se positionne plus uniquement en fournisseur de réseau pour transporter de la voix et des données, mais également en **fournisseur de contenu**, car il a l'avantage de pouvoir s'appuyer aujourd'hui sur une offre quadruple-play : téléphonie fixe, internet, télévision et téléphonie mobile.

France Telecom propose ainsi aussi bien :

- des films en Vidéo à la Demande (VOD),
- des chaînes de télévision (Orange Sport qui diffuse entre autres des matches de football de Ligue 1),
- les SMS  surtaxés (ex : concours, votes lors d'émissions de télé-réalité),
- les MMS ,
- les services mobiles sans contact, via les technologies NFC, RFID et SIM (par ex : permettre d'effectuer un paiement, valider un titre de transport, cumuler des points de fidélité... sur simple présentation d'un téléphone mobile à un lecteur sans contact),
- le Ring Back Tone : remplacer la tonalité d'attente liée à un numéro d'appel par un extrait musical ou sonore,
- la télévision mobile,
- la télévision mobile HD,
- la messagerie mobile,
- l'internet mobile,
- la navigation par GPS : Orange Maps,
- le chat : Orange Messenger (OWL),
- le wi-fi,
- la musique en streaming (partenariat avec Deezer)
- la messagerie vocale personnalisée,
- l'accès à internet sur ordinateur via la 3G : Internet Everywhere
- les téléchargements d'applications, sonneries, logos... : Application Shop,
- le partage de photos : Orange Photo...

La plupart des services multimédia sont hébergés sur la plate-forme de service MDSP  que je présenterai au chapitre suivant. Ci-dessous une représentation de l'accès à mon WebMail via la plate-forme MDSP  :

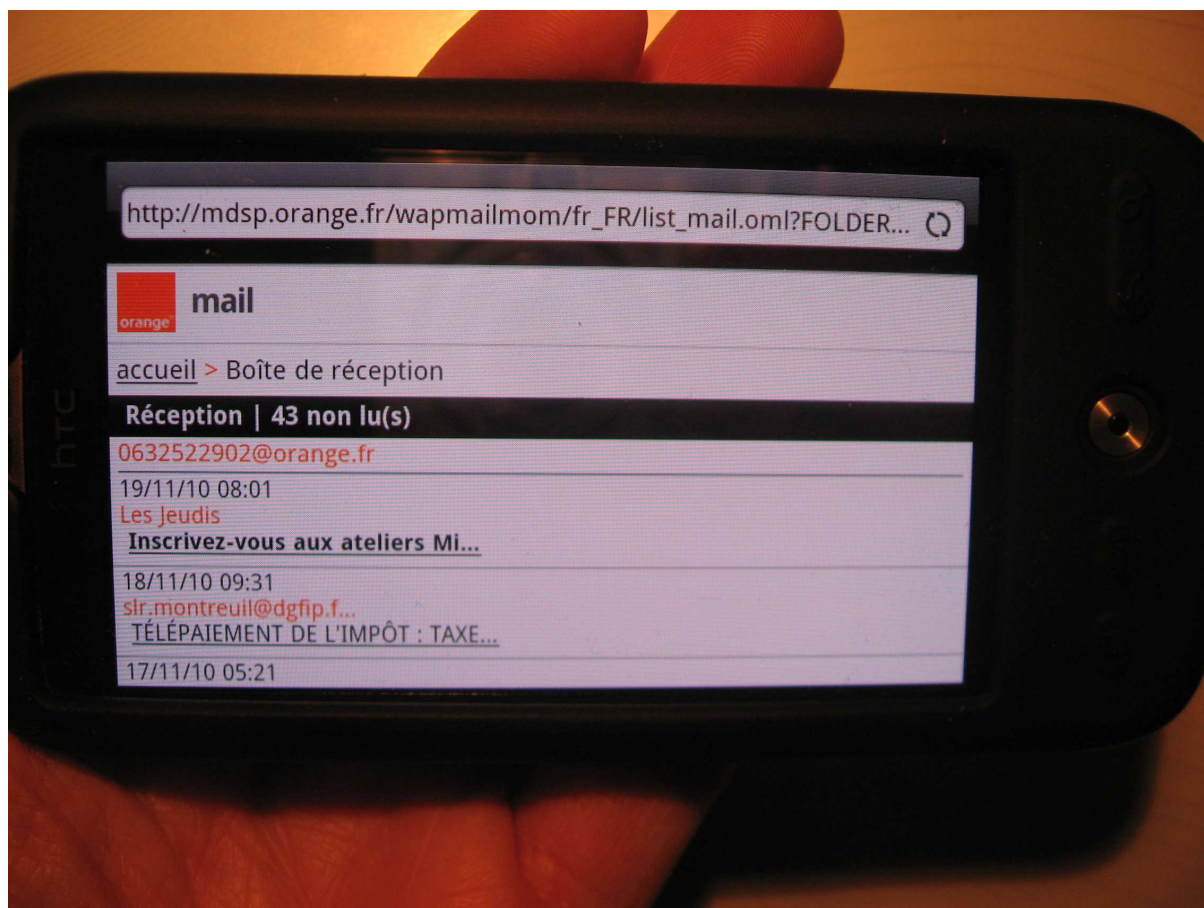


Illustration 3 : accès au mail Orange depuis son téléphone portable : via la plate-forme MDSP !

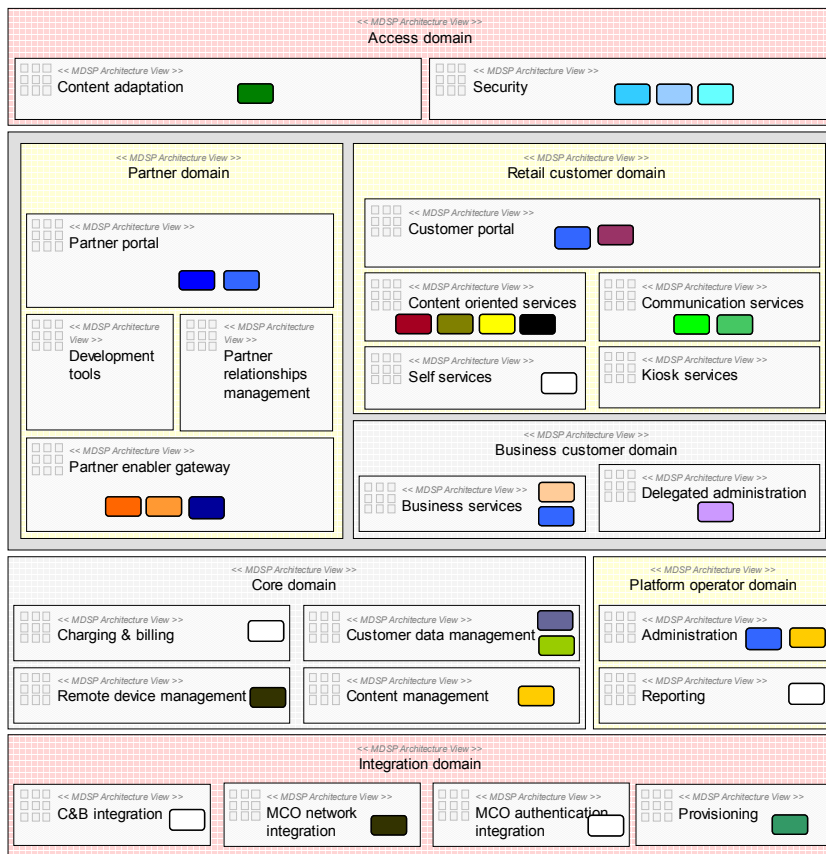
2.2. La plate-forme Mobile Data Service Platform (MDSP)

France Telecom a construit la plate-forme **Mobile Data Service Platform (MDSP)** dans l'ambition de **centraliser à l'échelle européenne tous les services de données mobiles Orange** sur une même plate-forme, et de mutualiser les solutions utilisées dans les différents pays européens pour réaliser des économies d'échelle.

Avec un potentiel de **55 millions de clients répartis dans 7 pays européens** (Belgique, France, Pays-Bas, Roumanie, Royaume-Uni, Slovaquie, Suisse), le projet MDSP se caractérise par la **complexité des développements** et par ses **nombreuses synergies** avec les projets transversaux du Groupe (New Datacenter, cloud, orientation green...).

MDSP est un projet transverse. Les services de données qui sont centralisés se composent, d'une part, de services de communication non conversationnels, comme le mail, le carnet d'adresse, le calendrier et la messagerie instantanée et, d'autre part, de services de contenus aussi variés que la météo, l'info, la bourse, les jeux, les logos/sonneries, le téléchargement d'une vidéo ou la recherche d'informations sur Internet.

La complexité des développements et de l'intégration d'un périmètre aussi large, et son planning serré, ont poussés France Telecom à lancer un appel d'offre. L'intégrateur **IBM Global Services** fut choisi en 2003 pour les phases de validation technique et déploiement massif. **Les logiciels utilisés** par l'intégrateur, proviennent d'une **vingtaine de fournisseurs différents**.



BESPOKE
MOBIXELL
INTERWOVEN
INTEGRO
AMBERPOINT
SYSTINET
ELATA
TSPACE
NETONOMY
MOTION BRIDGE
ORACLE
VOXMOBILI
VOX/WANADOO
VOLANTIS
WEBMETHOD
PACKETVIDEO
SWAPCOM
IBM EDGE SERVER
IBM Tivoli Access Manager
IBM Tivoli Federated Ident. Man.
IBM WebSphere Portal
IBM WES Appl. Portlet Builder
IBM WES Web Services Gateway
IBM WES Intell. Notif. Services
IBM WES User Privacy Manager

Illustration 4 : technologies présentes sur la plate-forme MDSP

Ci-dessous quelques services de la plate-forme MDSP :

- **OTAP** : Configuration GSM via SMS.
- **XMS-HUB** : Application d'envois SMS à d'autres applications.
- **OWL** (Orange Windows Live) : Messagerie instantanée sur mobile.
- **PIM** (Personal Information Management) : Flux de synchronisation des clients Orange des contacts/ agenda Mobile ou PC Personnel Information Management.
- **SIM IOD** : Solution dynamic SIM Tool Kit permet de gérer les menus après la délivrance de la carte sur le marché.
- **HHD-CHAH** : Home Hospital Discharge consiste en l'implémentation des ressources hospitalière dans le domicile des patients.
- **CSRTOOLS** : Agrégateur qui fournit une interface graphique multilingue pour l'administration Customer Services Support (CSR).
- **TEPEE** : Déclencher une alerte en cas de détection de fumée par appel téléphonique automatique.
- **Vending Machine** : Solution d'administration en temps réel d'un groupe de machine.

2.3. Le programme Multimedia Infrastructure Optimisation (MIO)

En 2007, France Telecom décide de reprendre la maîtrise d'œuvre et l'exploitation de la plateforme, assurées jusque là par l'intégrateur IBM. Dans cette optique, DPS lance le chantier **Multimedia Infrastructure Optimisation (MIO)** visant à optimiser l'ensemble de la plateforme MDSP📖.

La complexité du programme a conduit ses responsables à le diviser en **8 projets distincts** (hors projets récurrents) :

Projets	Début	Fin
Consolidated MIO Program - Global	24/09/2007	09/03/2009
Network Migration to IAS	24/09/2007	01/09/2008
Storage consolidation	24/09/2007	01/10/2008
Virtualisation	24/09/2007	01/10/2008
Public Zone Review	24/09/2007	24/12/2008
AIX Consolidation	24/09/2007	20/02/2009
Backup	24/09/2007	09/03/2009
Decommissionning	03/12/2007	24/12/2008
Tools Review	03/12/2007	24/12/2008

Tableau II : Les projets du programme MIO

Mon projet ayant été réalisé sur **le projet Backup**, la suite du document portera sur celui-ci. Cependant, j'ai eu des **interactions** à gérer avec les projets :

- **network migration to IAS** pour mutualiser les infrastructures IAS📖 et MDSP📖,
- **virtualisation** pour intégrer directement les machines virtualisées dans la nouvelle infrastructure de sauvegardes,
- **decommissionning** afin de s'assurer que les machines prises en compte dans la nouvelle infrastructure de sauvegarde ne fassent pas partie des machines décommissionnées,
- **AIX consolidation** pour intégrer les partitions AIX virtualisées dans la nouvelle infrastructure de sauvegarde.

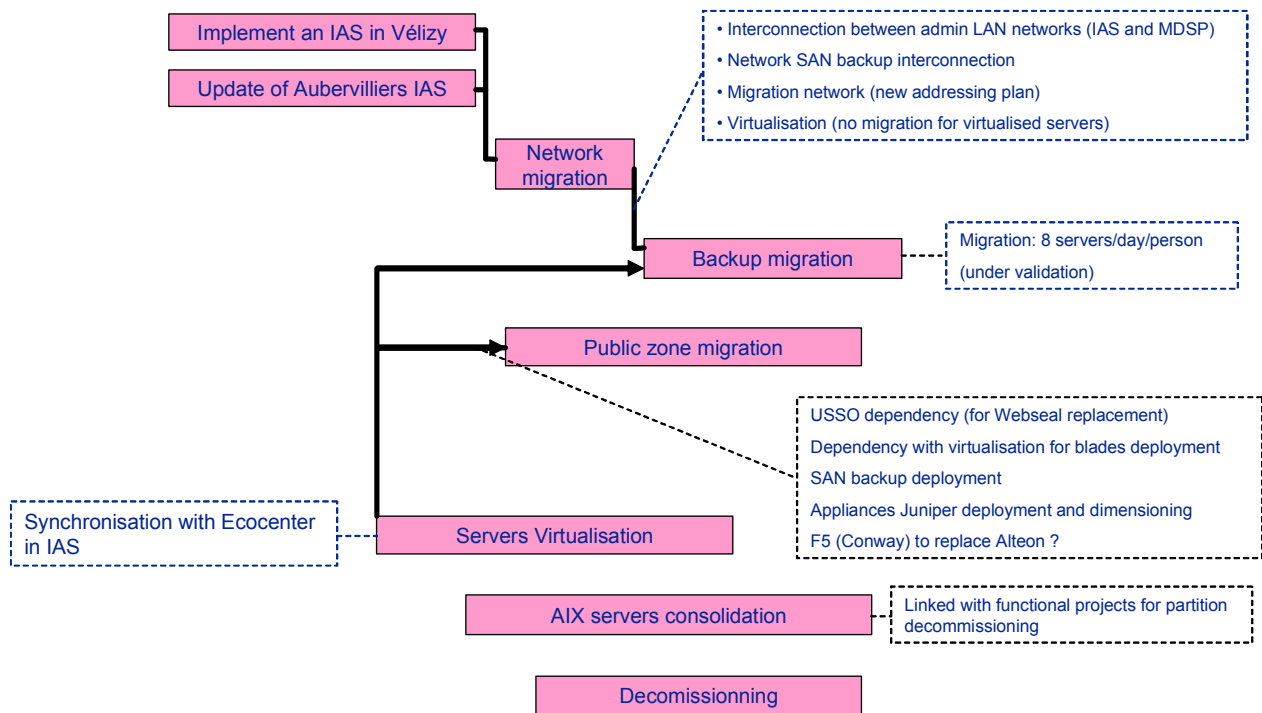


Illustration 5 : interactions entre les projets du programme MIO

2.4. Le projet MIO Backup

a) Objectifs

Le projet MIO Backup a pour objet **d'optimiser et de consolider l'infrastructure de sauvegarde** mise en place pour l'ensemble des serveurs de la plate-forme.

Dans le cadre du projet optimisation des sauvegardes, il a fallu répondre à trois besoins :

- **migration** du logiciel de sauvegarde (TiNa📖 vers NBU📖),
- **création d'une architecture mutualisée** avec les plates-formes de service en IAS📖 (pour plus de précision sur les **Infrastructure d'accès sécurisé**, se reporter à l'annexe I),
- **optimisation** des sauvegardes de base de données.

L'infrastructure de la plate-forme est **répartie sur 2 sites** :

- Aubervilliers hébergeant la production,
- Vélizy : hébergeant la plate-forme de **PRA**📖 d'Aubervilliers, et les pré-productions.

Cette répartition devant être conservée, les besoins furent différents selon le site :

- Aubervilliers : Étant également un site IAS📖, il a été nécessaire de faire une **évolution de l'architecture** NBU📖 existante,
- Vélizy : **Aucun site IAS**📖 n'étant présent, il a été décidé d'en **créer** un de toutes pièces.

Le projet MIO Backup englobe donc une **optimisation et une consolidation de l'infrastructure de sauvegarde** de la plate-forme MDSP, mais également la **mutualisation** de celle-ci avec les IAS. La nouvelle infrastructure devait être capable **d'absorber les besoins MDSP**, mais aussi être **suffisamment taillée** pour pouvoir accueillir les nouveaux projets des IAS de Vélizy et d'Aubervilliers.

À cet effet, les **objectifs et les axes d'optimisation** de la sauvegarde ont été définis comme suit :

- **migration de l'outil de sauvegarde** TiNa utilisé actuellement sur MDSP vers NBU, outil de référence France Telecom,
- **généralisation du mécanisme RMAN Lan Free** pour la sauvegarde des bases Oracle (à la place de RMAN Server Less et TiNa for Oracle),
- **généralisation de la sauvegarde sur disque**, c'est-à-dire sur Virtual Tape Library (VTL) pour remplacer les bibliothèques physiques actuellement utilisées pour les sauvegardes MDSP.

b) Enjeux de la solution de sauvegarde déployée

Consolidation : utilisation de la solution logicielle NBU déjà exploitée dans les IAS (sur le site d'Aubervilliers en particulier).

Optimisation de l'infrastructure de sauvegarde déployée : la solution VTL déployée sur les sites d'Aubervilliers et de Vélizy devra être mutualisable pour des demandes de sauvegardes des futurs services à héberger sur la plate-forme cible.

c) Contexte de la sauvegarde actuelle

L'infrastructure NBU est déjà **opérationnelle dans l'IAS d'Aubervilliers** : un master/media serveur mutualisé et une VTL DXi 5500 assurent déjà la sauvegarde des plate-formes de service de l'IAS d'Aubervilliers.

L'IAS de Vélizy reste à construire.

TiNa est le logiciel de sauvegarde actuel de MDSP mais la **fin du contrat de maintenance est fixée au 31/12/2008** et son renouvellement est déconseillé.

d) Cible

Déploiement ou évolution (suivant le site) de l'infrastructure de sauvegarde NBU dans les IAS sur les sites de Vélizy et Aubervilliers afin **d'accueillir** aussi bien **les migrations de la plate-forme MDSP** que toute nouvelle **plate-forme de service MDSP dont le service doit être ouvert en 2009**.

À fortiori, l'IAS de Vélizy doit ouvrir avec un service de sauvegarde opérationnel pour **accueillir** toute **plate-forme de service prévue** sur le **Datacenter de Vélizy**.

e) Qualité de service

Un objectif non écrit mais évident est bien entendu la conservation voire l'amélioration de la qualité de service qui se traduit par :

- le taux de réussite des sauvegardes (on observe jusqu'à 20% d'échec avec les sauvegardes TiNa actuelles, dues essentiellement aux erreurs physiques des bandes ou des lecteurs, et à la complexité du processus de sauvegarde des bases de données),
- le taux de disponibilité de l'infrastructure de sauvegarde,
- les performances en sauvegarde et en restauration (on observe plusieurs heures d'attente de ressources lecteurs).

2.5. Périmètre initial


Je ne présenterai ici que les hypothèses en début de projet ; je présenterai dans le chapitre 5 le résultat des audits que j'ai menés sur des points particuliers.

Le périmètre initial du projet comprenait environ **1200 machines pour la plate-forme MDSP** (environnements de pré-production et production) sauvegardées avec l'outil de sauvegarde TiNa en version 3.7 sur des robotiques physiques.

En prévision des premiers décommissionnements à venir, le périmètre MDSP a été **restreint à environ 1000 machines** à migrer.

Il faut cependant souligner la complexité de certains serveurs qui **hébergent une base de données de forte volumétrie en cluster sauvegardée en LAN Free** (via le SAN).




3. L'ÉTAT DE L'ART

Je vais présenter dans ce chapitre la sauvegarde telle qu'elle est préconisée par mon équipe (MOE  Sauvegarde) et donc implémentée dans toutes les entités et filiales du groupe France Telecom.

3.1. Le but de la sauvegarde

La sauvegarde a pour but de protéger les données de tout type de corruption (erreur humaine, panne matérielle...) afin de pouvoir les restituer. Le défi est de pouvoir sauvegarder les données critiques de manière cohérente, le plus rapidement possible, en gênant le moins possible les utilisateurs et les traitements, et de garantir une restauration rapide d'un état du serveur à un moment donné.

3.2. Les principes de la sauvegarde

Le fonctionnement de la sauvegarde est simple : les serveurs applicatifs envoient leurs données depuis le stockage primaire (disques locaux, baie SAN  ou filer NAS ) vers un stockage secondaire (une robotique de sauvegarde) et envoient leur méta-données , qui indexent des informations relatives à la nature et la localisation des données, dans un catalogue.

Les sauvegardes peuvent emprunter 2 chemins :

Pour des volumétries de sauvegarde faibles (moins de 50 Go) et des besoins de débit faibles, les données transitent depuis le serveur applicatif jusqu'à un serveur de sauvegarde via le réseau IP ; on parlera de sauvegarde **LAN**.

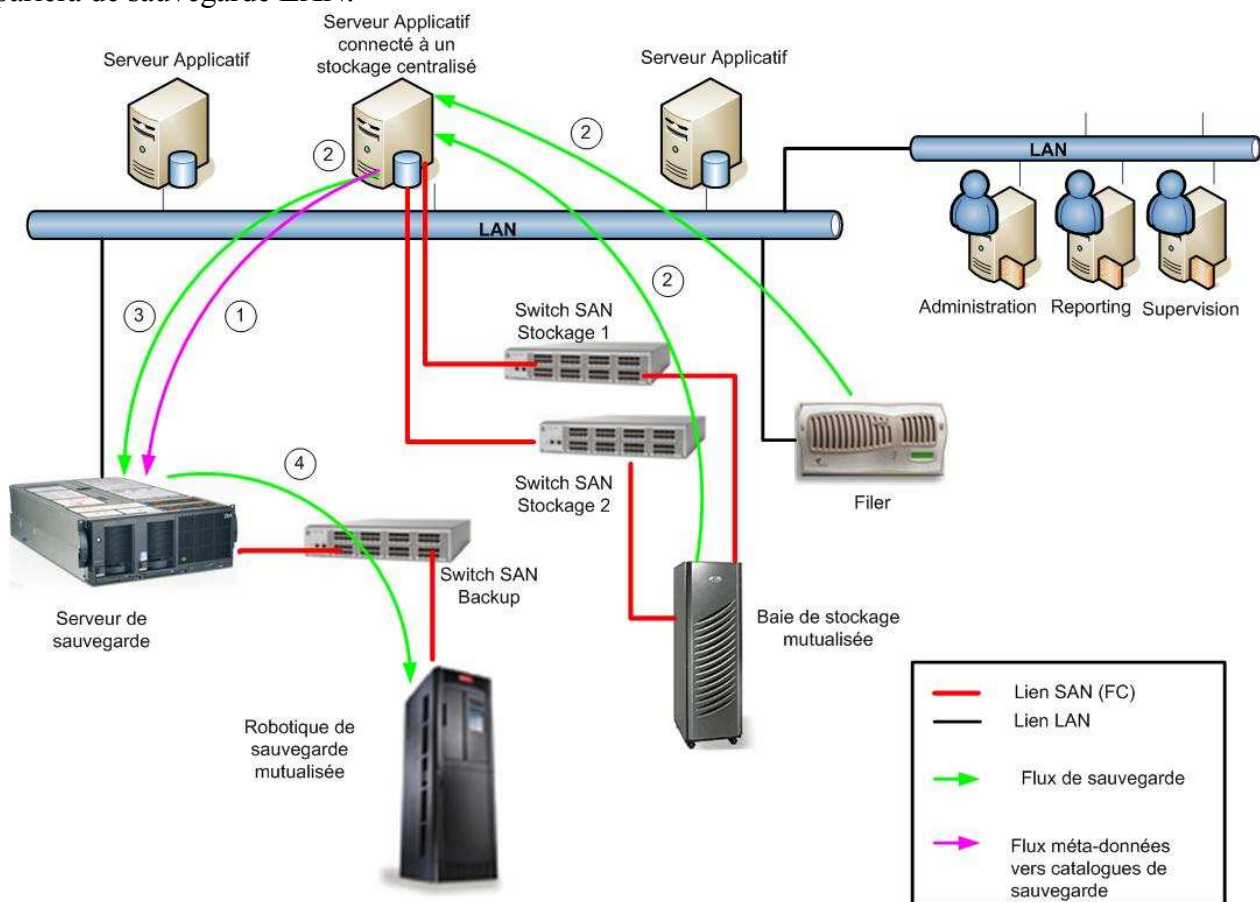


Illustration 6 : sauvegarde LAN

Les méta-données📖 sont envoyées depuis le serveur applicatif vers le serveur hébergeant le catalogue (*flux 1*). Les données à sauvegarder sont extraites par le serveur applicatif depuis les disques locaux, une baie SAN📖 ou un filer NAS📖 (*flux 2*), et envoyées vers un serveur de sauvegarde (*flux 3*) lequel accède à la robotique via le SAN📖 (*flux 4*) et se charge de la mise sur bandes.

À l'inverse, dans le cas de volumétries plus importantes ou des besoins de débit importants, le serveur applicatif est raccordé au réseau SAN pour accéder à la robotique. On parlera de sauvegarde **SAN** ou **LAN Free**.

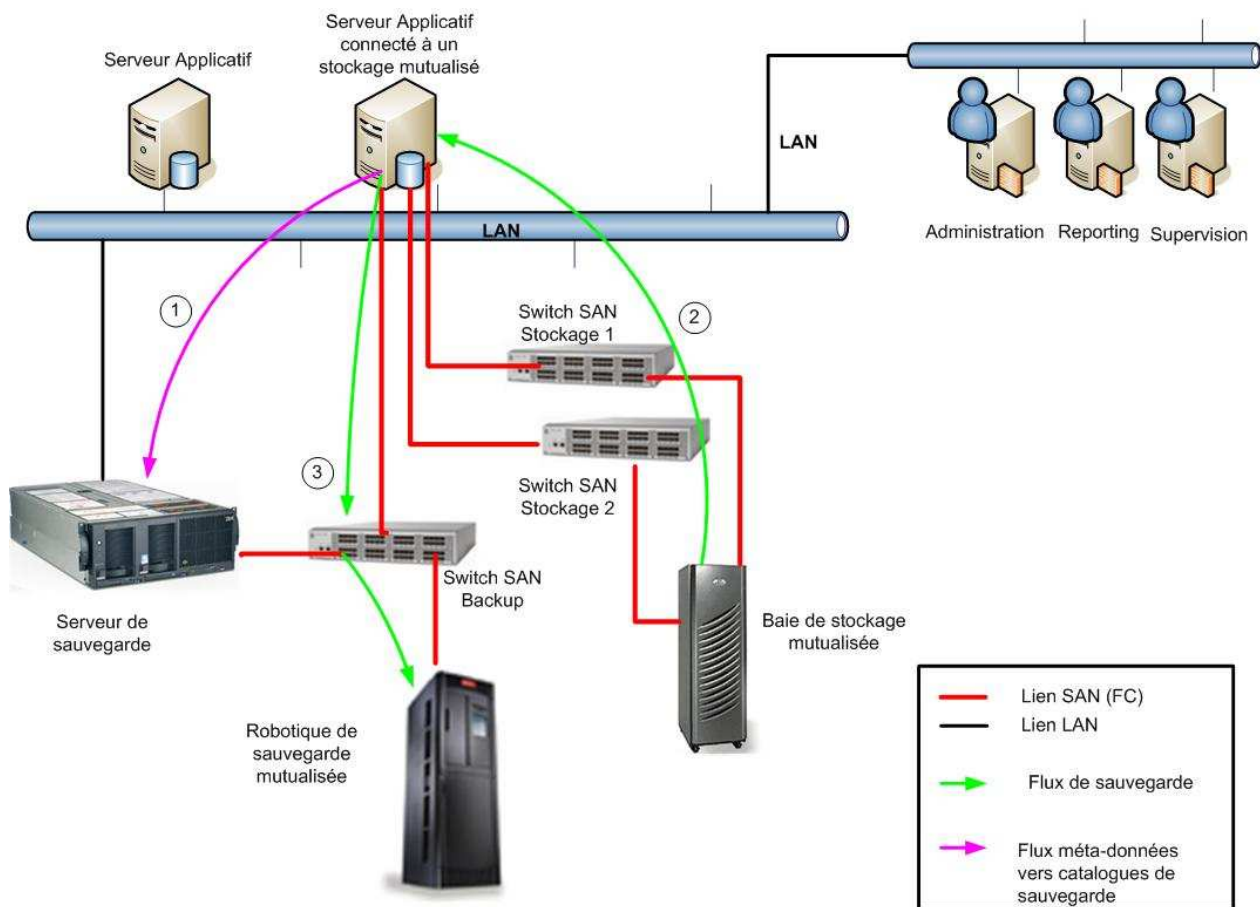


Illustration 7 : sauvegarde SAN ou LAN Free

Les méta-données sont envoyées depuis le serveur applicatif vers le serveur hébergeant le catalogue (*flux 1*). Les données à sauvegarder sont extraites par le serveur applicatif depuis les disques locaux, une baie SAN ou un filer NAS (*flux 2*) ; le serveur applicatif accède à la robotique via le SAN (*flux 3*) et se charge de la mise sur bandes.

On peut distinguer le cas particulier des sauvegardes NDMP des filers NAS. En effet on peut sauvegarder un filer NAS comme présenté dans les 2 cas ci-dessus c'est-à-dire via le serveur applicatif, mais avec l'inconvénient que les données transitent via le LAN entre le serveur applicatif et le filer NAS. Il est souvent préférable de sauvegarder directement les volumes du filer via le SAN grâce au protocole NDMP.

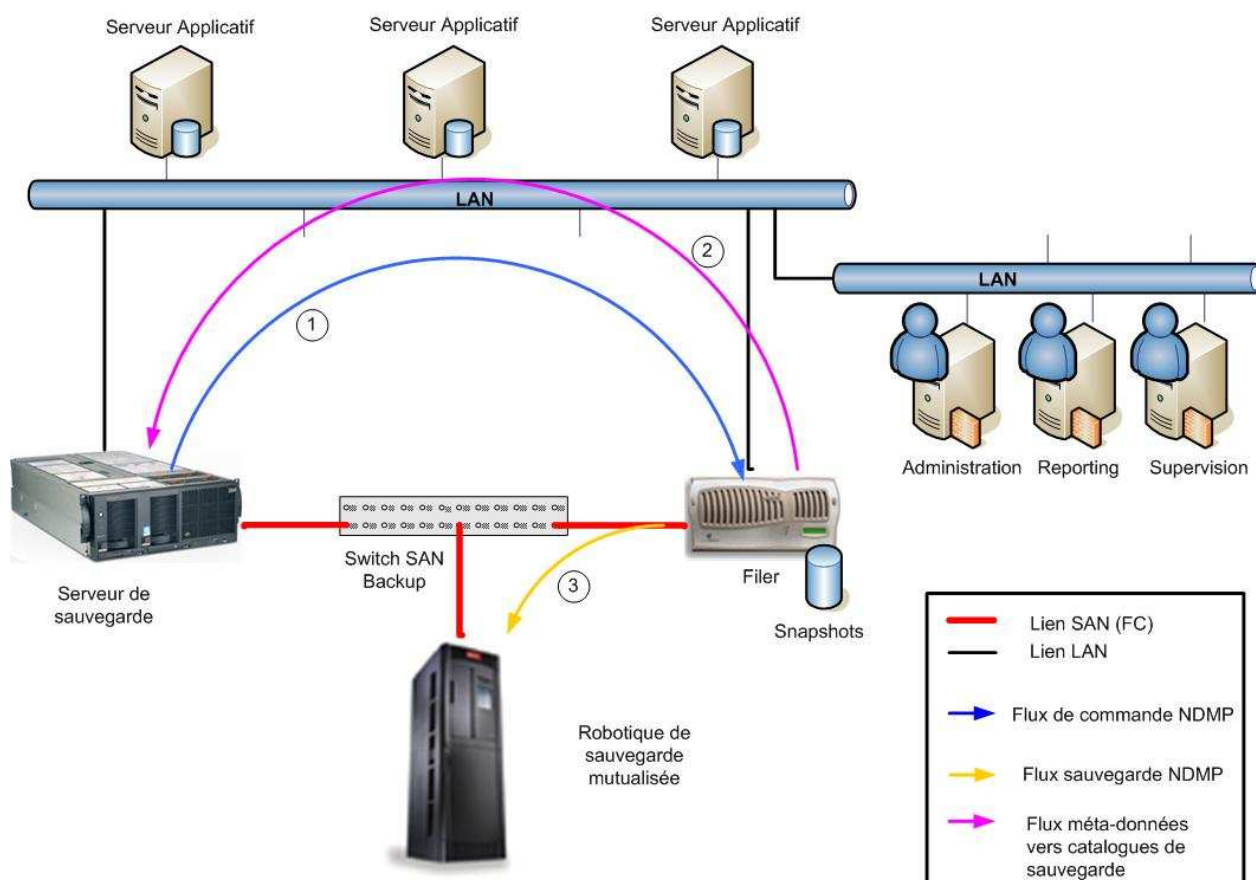


Illustration 8 : sauvegarde NDMP

Le serveur de sauvegarde envoie des commandes au filer afin qu'il génère un snapshot puis envoie ses données à la robotique (*flux 1*). Les méta-données sont envoyées depuis le filer NAS vers le serveur hébergeant le catalogue (*flux 2*). Le filer accède à la robotique via le SAN (*flux 3*) et se charge de la mise sur bandes des données à sauvegarder.

On distingue généralement 4 types de sauvegardes :

- **sauvegarde «Applicative»** pour une reprise en cas d'incident applicatif
- **sauvegarde «Système»** pour une reprise en cas d'incident sur le système
- **sauvegarde «Site»** pour :
 - o une reprise en cas de destruction du site ou de la plate-forme
 - o une reconstruction de la plate-forme en cas de déménagement
- **sauvegarde «Externalisée»** pour permettre :
 - o une restauration sur un site distant
 - o une injection de données de production en pré-production

Et 3 méthodes :

- **sauvegarde totale (ou full)**
 - ne prend pas en compte les sauvegardes précédemment effectuées
 - ne sauvegarde pas pour autant tous les fichiers
- **sauvegarde incrémentale**
 - ne sauvegarde que les fichiers modifiés depuis une sauvegarde de référence précédemment effectuée
 - la terminologie diffère selon les logiciels de sauvegarde. Par exemple pour NetBackup, une incrémentale cumulative se base sur la dernière sauvegarde totale, alors qu'une incrémentale différentielle se base sur la dernière sauvegarde incrémentale (différentielle ou cumulative)
- **sauvegarde synthétique**
 - sauvegarde totale reconstituée au niveau du serveur de sauvegarde sur la base de la totale de référence et des incrémentales qui ont suivi (dispense de reprendre l'intégralité des données à la source).


Selon le temps d'indisponibilité autorisé des données primaires, la sauvegarde peut s'effectuer :

- **à chaud** : on sauvegarde l'applicatif alors qu'il est accédé et/ou modifié. Ceci peut se faire :
 - via un agent ou module du logiciel de sauvegarde qui s'interface avec l'applicatif ; il existe ainsi des agents NBU et TiNa pour Oracle📖, DB2📖, Exchange📖...
 - via un snapshot📖.
- **à froid** : on arrête l'applicatif avant la sauvegarde, puis on le redémarre après la sauvegarde. Le logiciel de sauvegarde peut gérer cet ordonnancement via des scripts de pré- et post-traitement.
- **tiède** : on «gèle» l'applicatif le temps de recopier les données sur un espace tampon, puis on redémarre l'applicatif. Les données sont sauvegardées via un serveur tiers (appelé parfois «Infocentre») qui accède à l'espace tampon.

Les sauvegardes doivent répondre à des contraintes :

- **la durée inférieure à 24H pour le cycle applicatif journalier**
 - pour effectuer la sauvegarde
 - mais aussi les traitements utilisateurs
- **minimiser les perturbations applicatives**
 - utilisation des plages horaires creuses
- **les performances**
 - un niveau de performance minimum doit être garanti à l'utilisateur final

- l'intégrité

- il est inutile de sauvegarder des données inutilisables ou incohérentes (par exemple : sauvegarder les datafiles d'une base de données Oracle  en cours de fonctionnement).

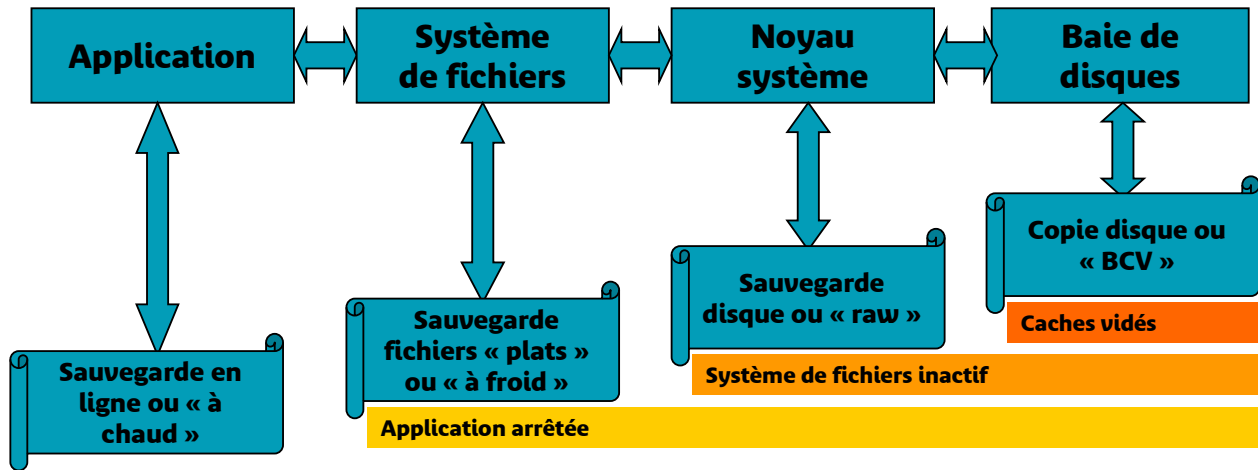











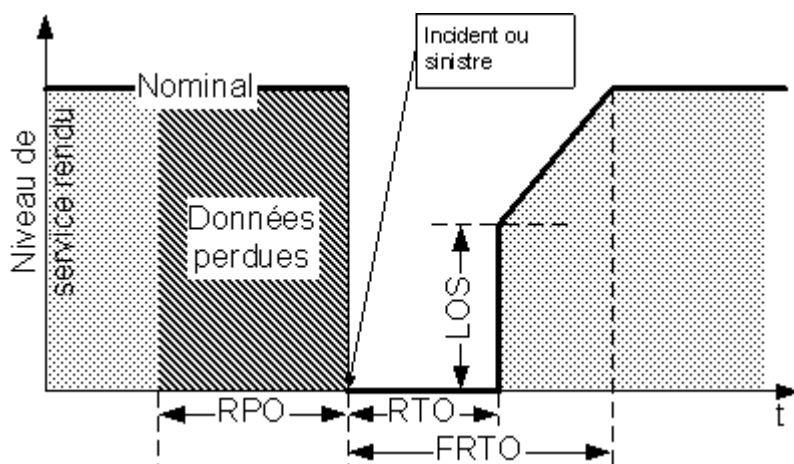
Illustration 9 : contraintes pour garantir l'intégrité des données



Dans le cas d'une application répartie sur plusieurs serveurs, aucun mécanisme générique validé ne permet d'assurer la cohérence globale de la sauvegarde. Cette cohérence des données sauvegardées doit reposer sur des mécanismes fournis par chaque projet.

Les sauvegardes sont essentielles dans le cadre des **plans de secours** :

- **contexte: reprise d'activité après sinistre localisé ou global**
 - PCA : Plan de Continuité d'Activités
 - PRA  : Plan de Reprise d'Activités
 - PRAp: Plan de Reprise Applicative
 - PRI : Plan de Reprise Infrastructure
 - SLA  : Service Layer Agreement → Contrat spécifiant les niveaux de services attendus
- **par application, deux engagements majeurs de l'exploitant:**
 - RTO  : durée maximale d'indisponibilité totale d'une application
 - RPO  : perte de donnée maximale admissible
- **impact sur le Recovery Point Objective (RPO ) → Une sauvegarde par jour**
 - la sauvegarde ne permet pas d'assurer un PRA  avec $RPO  < 1J$
 - on peut cependant permettre la sauvegarde d'archives logs à une fréquence supérieure, en gardant à l'esprit que la sauvegarde du catalogue de l'outil de sauvegarde peut être réalisée à une fréquence différente.
- **impact sur le Recovery Time Objective (RTO ) → Temps de restauration des données**
 - le temps de restauration est sur le chemin critique du RTO 



RTO 📖 : Recovery time objective : Durée maximale d'indisponibilité totale

FRTO : Full RTO : RTO nominal

LOS : Level Of Service : Puissance de traitement minimale garantie en mode dégradé 📖

RPO 📖 : Recovery point objective : Perte maximale de données

Illustration 10 : RPO et RTO



Il est important de rappeler que le but de toute sauvegarde reste de pouvoir restaurer les données suivant les contraintes de RTO 📖 et de RPO 📖. Ainsi, l'élaboration de la stratégie de sauvegarde adéquate doit être réalisée dans ce but. Néanmoins lorsque le RTO 📖 et/ou le RPO 📖 sont inférieurs à 24h, ceux-ci ne devront pas être garantis par la sauvegarde et ce quelque soit la volumétrie. D'autres mécanismes devront alors assurer le RPO 📖 et le RTO 📖 (réplication de baies par exemple) ; ces répliques ne peuvent cependant être considérées comme de la sauvegarde (pas d'historisation).

Pour le SI 📖, le plan de secours se décline en trois niveaux :

- Le Plan de Continuité des Activités (**PCA**) ou Business Continuity Plan (BCP) représente l'ensemble des mesures prises à l'avance pour faire face à un sinistre majeur qui mettrait hors service tout ou partie des moyens vitaux de l'entreprise. Il donne le schéma complet de reprise d'activité d'une entreprise via des procédures : processus métiers, ressources humaines, bâtiments, systèmes informatiques et de télécommunications, informations clients...
- Le Plan de Reprise après Sinistre (**PRS**) ou Disaster Recovery Plan (DRP), au sens processus métiers, représente les actions à prendre (SI 📖, mais aussi bâtiments, localisation des opérateurs...) pour reprendre un service métier et dans quelles conditions.
- Le Plan de Reprise Applicatif (**PRAp**) ou Application Disaster Recovery (ADR) est la reprise d'une application selon les contraintes applicatives qui découlent du PRS. Ce dernier détermine les besoins et contraintes de la **Continuité de Service** du SI 📖 et notamment conclura sur le dimensionnement et les besoins en **Haute Disponibilité** pour chaque application.

3.3. L'évolution de la sauvegarde

Deux grandes problématiques, liées à l'externalisation des données et à l'utilisation simultanée des lecteurs, sont à l'origine des différents projets de standardisation de l'infrastructure de sauvegarde.

Les lecteurs de bandes sont en effet chargés d'absorber les flux de données dans des plages horaires très courtes (en général la nuit). Toutes les applications sont sauvegardées en même temps, nous sommes alors obligés de multiplier le nombre de lecteurs de bandes en entrée. Cela induit des coûts conséquents car les lecteurs de bandes représentent les investissements les plus lourds dans l'infrastructure de sauvegarde.

Parallèlement, pour assurer une sécurité maximale, il est fortement préconisé d'externaliser les données à sauvegarder. Toutefois, des contraintes liées aux distances et à la performance des systèmes empêchent de faire transiter des données volumineuses d'un site à un autre.

Aujourd'hui, cette externalisation est effectuée manuellement par les techniciens de proximité, qui chaque jour récupèrent les bandes dans la robotique et les emmènent sur un autre site. Cette procédure comporte plusieurs inconvénients : risque d'erreur humaine liée à la gestion du référentiel des bandes, risque de perte des bandes, perte de temps...

La solution de virtualisation des lecteurs de bandes peut répondre aujourd'hui à ces besoins.

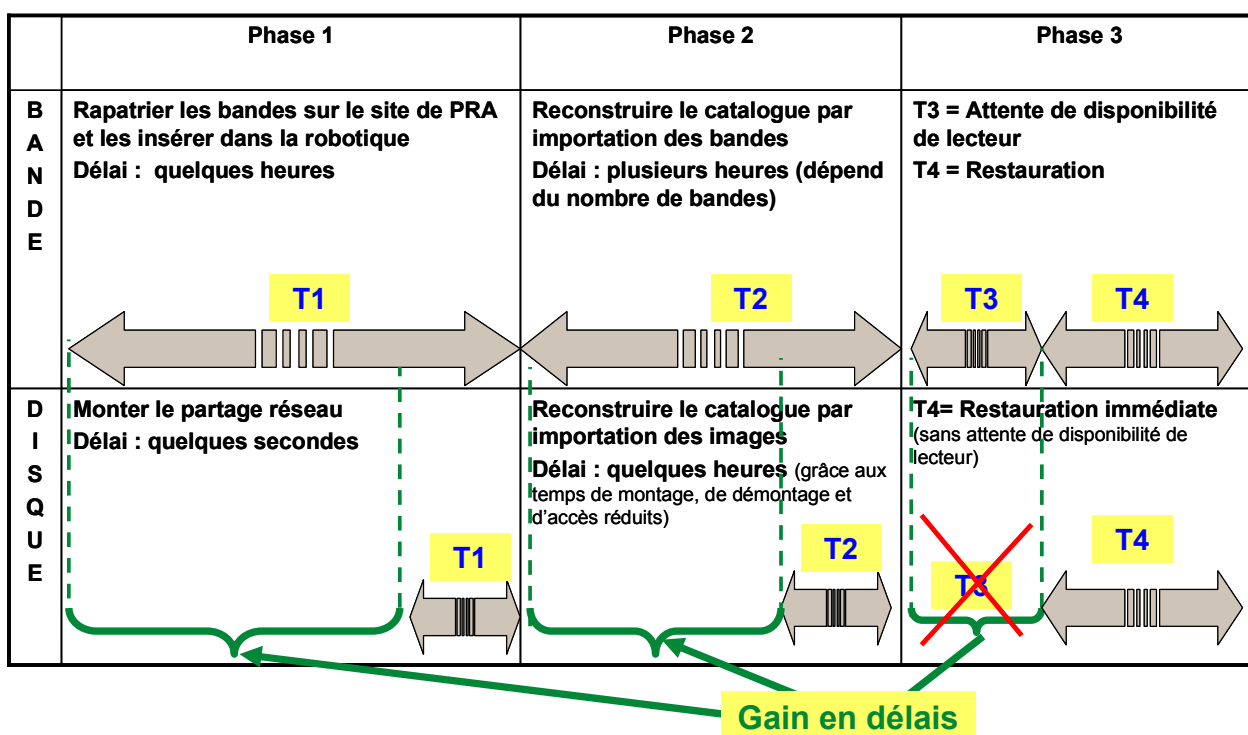


Illustration 11 : gains sur les processus de reprise

3.4. La virtualisation

Les utilisateurs restreignent souvent la seule virtualisation à celle des serveurs. Il faut cependant **distinguer les solutions** de :

- **virtualisation de serveurs Unix** (la solution dans ce cas étant propriétaire et spécifique à chaque constructeur) ou **x86** (Windows ou Linux).
- **virtualisation de stockage** permettant de faire migrer un espace de stockage d'une solution coûteuse sur SAN/FC vers du iSCSI sur disques SATA ou inversement suivant le besoin à un instant T et ce en toute transparence au niveau du serveur et de l'application.
- **virtualisation de sauvegarde** utilisant une VTL émulant une robotique bandes, mais dont le stockage est en réalité réalisé sur disques, ce qui permet de plus de bénéficier sur option de la déduplication. *La déduplication sera présentée en détail au chapitre 3.5.*
- (à venir) **virtualisation du réseau** permettant de déporter la fonction actuellement réalisée par les systèmes de virtualisation des serveurs au niveau des équipements réseau.

Mon mémoire portant bien entendu sur la sauvegarde, je parlerai essentiellement de la virtualisation de la sauvegarde sur la base de deux technologies issues des fournisseurs Fujitsu-Siemens et Quantum.

La solution de Fujitsu-Siemens, le CentricStor, consiste à sauvegarder les données sur le site nominal dans une VTL tampon puis à les répliquer via FC dans une robotique physique sur un site distant.

Cette solution comporte au moins deux atouts. D'une part, le nombre de lecteurs de bandes est divisé par 5 car le cache-disque alimente au fur et à mesure les lecteurs de bandes (alors qu'ils n'étaient exploités qu'à 10% de leur capacité auparavant). D'autre part, les restaurations de données les plus courantes (J - 3) sont réalisées plus rapidement quand les données sont stockées sur le cache-disque.

Cette infrastructure de virtualisation se trouve actuellement déployée sur tous les Datacenters Parisiens de la Direction des Opérations du SI (DOSI) où j'exerçais précédemment.

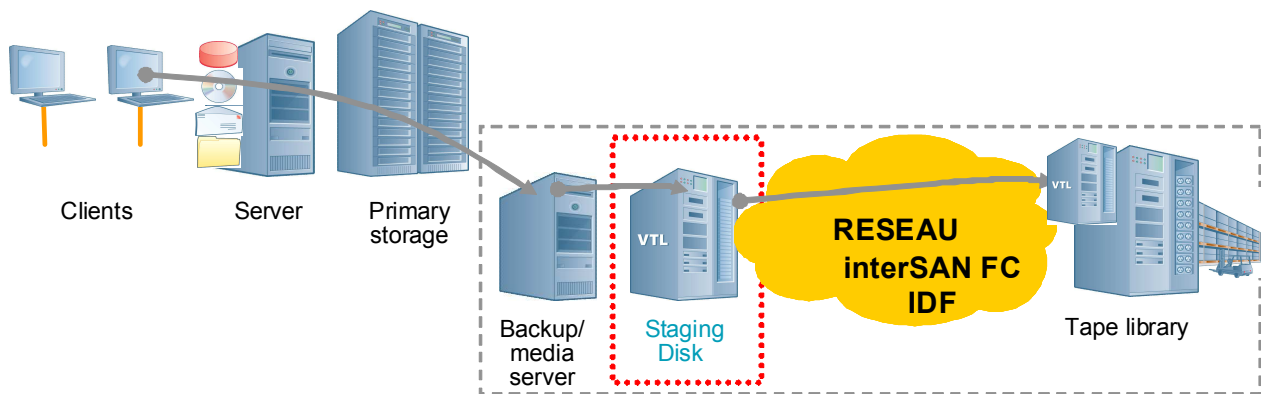


Illustration 12 : principe d'implantation du CentricStor

La solution concurrente de Quantum, le DXi, consiste à sauvegarder les données sur le site nominal dans une VTL. On peut dans un second temps les répliquer via IP dans une autre VTL sur un site distant.

Cette infrastructure de virtualisation se trouve actuellement déployée sur tous les IAS de la Direction des Plates-formes de Service (DPS) où j'exerce actuellement.

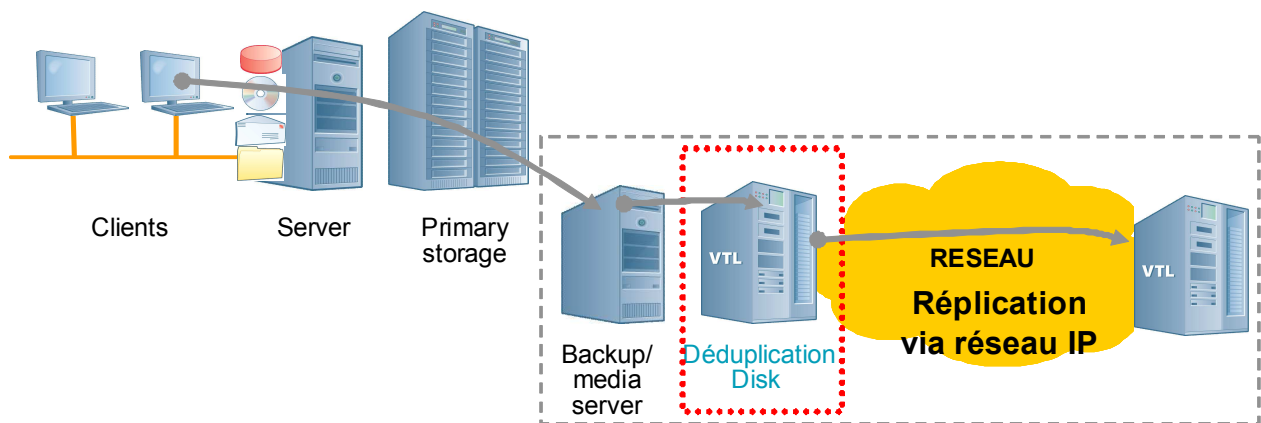


Illustration 13 : principe d'implantation du DXi

La consolidation de l'infrastructure de sauvegarde constitue un projet complémentaire qui consiste à mettre en œuvre des robotiques et des lecteurs de consolidation plus performants et plus capacitifs dans un processus standardisé.

3.5. La déduplication

La déduplication^[1], également appelée factorisation ou stockage d'instance unique, consiste à optimiser l'espace de sauvegarde en détectant les segments de données redondants. Les fichiers sont découpés en une multitude de tronçons, auxquels est associé un identifiant unique. La comparaison de ces identifiants permet de ne stocker qu'une seule fois un même tronçon. Les données redondées sont remplacées par des pointeurs stockés dans un index. Cette factorisation des séquences de données identiques permet d'économiser drastiquement l'espace de stockage.

Les principaux bénéfices sont une réduction de la taille des données :

- stockées et donc un gain d'espace et d'énergie (meilleure densité),
- en transit et donc une réduction de la bande passante nécessaire,
- et donc au total un gain financier..

Tous les acteurs de la sauvegarde se positionnent sur cette technologie : Diligent, Avamar, Data Domain, Sepaton pour les éditeurs de niche ; et Symantec, EMC, HP, HDS, Adic/Quantum, Overland, et Falconstor pour les plus grands noms.

a) Le principe

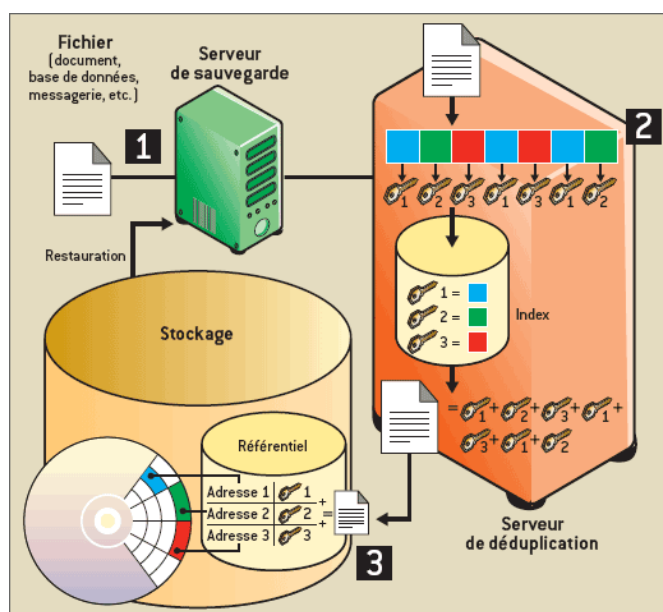



Illustration 14 : principe de la déduplication

Un processus centralisé ou décentralisé :

Certains serveurs de déduplication^[1] (Adic/Quantum et Data Domain, par exemple) s'insèrent dans le flux des données de sauvegarde, souvent en émulant une robotique virtuelle (VTL^[1]). D'autres systèmes (notamment PureDisk ou Avamar) sont plus décentralisés : les processus de déduplication^[1] sont répartis sur les serveurs de production au moyen d'agents. Les données sont ensuite envoyées à un serveur de déduplication^[1], qui, dans cette architecture, se substitue au serveur de sauvegarde.


Une segmentation déterministe ou aléatoire :

Le fichier envoyé par le serveur de sauvegarde est découpé par l'algorithme de déduplication  en plusieurs blocs. À chaque portion est associée une signature unique, sous la forme de «hash». Selon les technologies, le découpage est soit déterministe (un segment de 8 Ko, par exemple), soit aléatoire (sa taille diffère selon les séquences de bits identifiées par l'algorithme). Tous ces «hashs» sont ensuite comparés à ceux déjà stockés dans l'index. Lorsqu'un «hash» est présent, c'est que le segment de données qui lui est associé est déjà stocké. Si aucune correspondance n'est trouvée, la séquence est stockée, et l'index mis à jour.

De l'index aux adresses physiques :

L'index envoie au référentiel la description du fichier à stocker sous forme d'une combinaison de signatures. Le référentiel traduit alors ces signatures en pointeurs vers les adresses physiques des séquences de données. Lorsque le serveur de sauvegarde restaure des données, il s'adresse directement à ce référentiel.

b) L'algorithme

Le principal enjeu consiste à trouver le meilleur équilibre entre la factorisation des séquences (la plus élevée possible) et la taille (la plus réduite possible) de l'index des signatures. Plus les tronçons sont petits, plus la factorisation - et donc l'économie d'espace de stockage - devient importante. Inversement, les signatures générées sont plus nombreuses, et alourdissent l'index. Aussi, pour des questions de performance, celui-ci doit être stocké dans le cache du serveur de déduplication , et non sur disque. Et cette mémoire cache s'avère limitée et coûteuse.

Sur le papier, Diligent d'IBM affiche, à ce jour, les meilleures performances. Avec 4 Go de cache, il adresse 1 Po (1 petaoctet, soit 1 000 téraoctets) de données, à un débit moyen de 200 Mo/s (selon le contenu des fichiers). Son secret réside dans un algorithme de hachage propriétaire, et non sur SHA-1 ou MD5 pour signer les tronçons. Il optimise ainsi la taille des segments en fonction des séquences de bits identifiées.

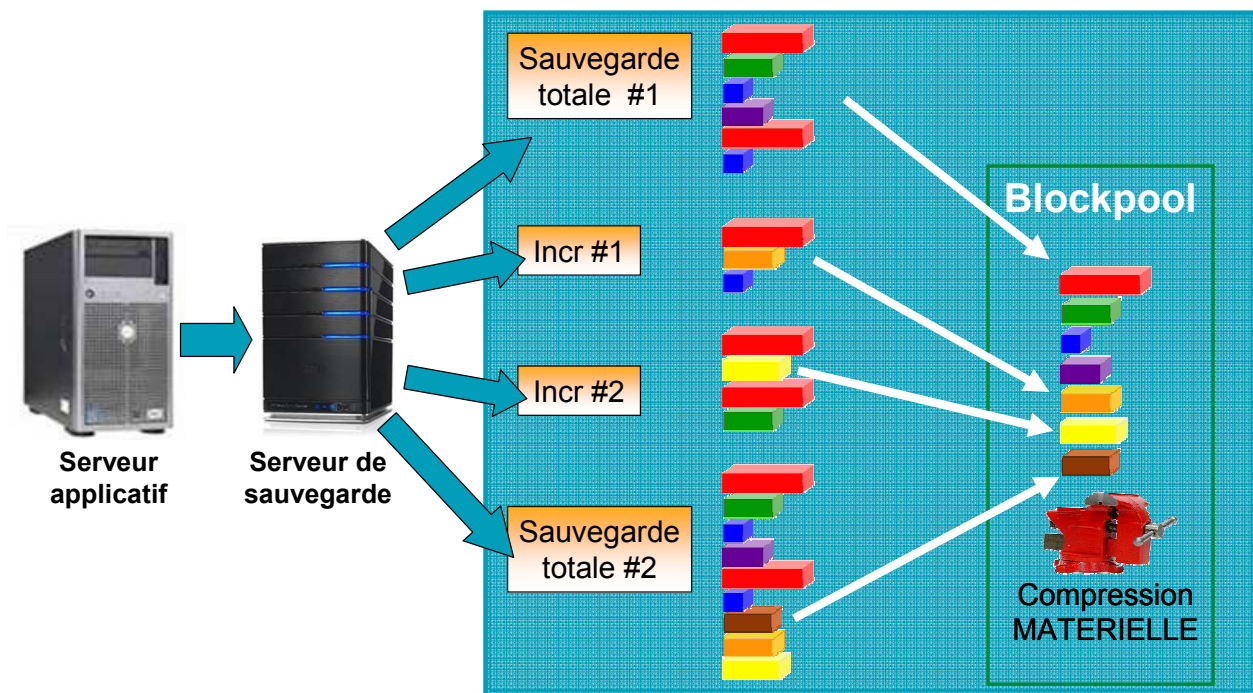


Illustration 15 : exemple de déduplication à la cible

c) Évolution vers le stockage primaire

Le constructeur NetApp embarque cette technologie dans ses baies de production FAS. Selon un responsable marketing de la société NetApp (spécialisée dans les solutions NAS) "La déduplication est une technologie stratégique pour faire face à l'explosion des volumes de stockage dans les entreprises. Aujourd'hui et demain, le meilleur fournisseur de stockage ne sera pas celui qui vendra les plus gros volumes, mais celui qui permettra d'en économiser le plus". C'est pourquoi Netapp a intégré son propre moteur de déduplication baptisé A-SIS au sein de ses gammes de baies FAS (SAN et NAS).

NetApp est le premier constructeur à intégrer directement cette technologie dans un équipement de stockage et non sur un serveur dédié, et donc à l'exploiter sur une baie de stockage primaire (et non dans un processus de sauvegarde).

Dans ce cas, les performances atteintes sont moindres que lors d'une exploitation en sauvegarde et l'on peut s'attendre à une réduction de l'ordre de 1,5 à 2,5 du volume des données mais ce gain dépend essentiellement de l'application.

d) Adoptée par toute l'industrie du stockage

Tous les géants du stockage se sont positionnés sur cette technologie comme en témoignent les rachats de Datacenter Technologies par Symantec, de Rocksoft par Adic/Quantum ou d'Archivas par Hitachi Data Storage. Selon l'éditeur EMC qui a mis la main sur Avamar, ce dernier permet une réduction des temps de sauvegarde des machines virtuelles pouvant atteindre 90 %. Il y a fort à parier que dans quelques temps, plus aucune baie de disques ne se vendra sans moteur de déduplication intégré.

3.6. Conclusion

Afin de protéger les données du projet MDSP, je vais utiliser les techniques ci-dessous :

Le rattachement au SAN impliquant un surcoût (carte HBA, câblage FC, affectation de ports sur le switch SAN, attribution de lecteur de bande...), cette solution ne peut être généralisée pour tous les serveurs de la plate-forme. Je vais **privilégier les sauvegardes SAN pour les serveurs à forte volumétrie** et/ou qui demandent un débit important. Cette méthode sera donc essentiellement mise en place pour les sauvegardes des bases de données. Le fait de déplacer ces sauvegardes sur le SAN permettra également de réduire la charge d'utilisation du LAN pour les autres serveurs. Les données de la plate-forme MDSP étant stockées sur les disques locaux des serveurs et sur des baies de stockage SAN, et non sur des filers NAS, je ne mettrai pas en place de sauvegarde NDMP sur ce périmètre.

Les sauvegardes incrémentales permettent de sauvegarder uniquement le delta des données modifiées, et donc de réduire le volume de la sauvegarde. En cas de restauration, il faut par contre restaurer une sauvegarde totale ou synthétique et une ou plusieurs incrémentales, lesquelles peuvent être réparties sur un grand nombre de bandes. Afin de limiter le temps de restauration, il faut donc refaire une sauvegarde totale ou synthétique régulièrement. La sauvegarde synthétique permet de moins solliciter le client et le réseau, mais ajoute de la charge sur le serveur (Media Serveur qui sera largement présenté plus tard) qui fait le traitement. Ce type de sauvegarde pourra être intéressant ponctuellement pour des clients au débit réseau faible.

Je m'assurerai que les sauvegardes peuvent se dérouler intégralement dans les plages horaires utilisées. La mise en place d'une Boucle Qualité Nationale (BQN) permettra de suivre les échecs et les performances de la solution déployée. La validation des EB permettra de s'assurer que l'on ne sauvegarde pas des données inutiles ou non intègres.

Via, entre autres, l'externalisation des données, la solution déployée devra permettre d'assurer le PRA des applications, dans le respect du RPO et du RTO demandés par la MOA. Les nouvelles solutions de virtualisation de la sauvegarde et de la déduplication seront étudiées dans le cadre de ce projet, et pourront être implantées en cas d'apport technique et d'intérêt économique.

4. ORGANISATION DU PROJET

4.1. Méthodologie de la conduite du projet

Pour qu'il soit géré dans un contexte de qualité, un projet doit suivre différentes phases au terme desquelles des points de contrôle doivent être validés. Chaque étape fait l'objet d'un livrable et d'une validation à partir d'un document spécifique. Cela permet de maîtriser la conformité des livrables à la définition des besoins ainsi que de s'assurer de l'adéquation aux objectifs de coûts et de délai.

La méthodologie de conduite de projet en vigueur chez France Telecom est le processus Time To Market (**TTM**). Le TTM est l'ensemble des actions génériques, jalons, livrables et règles de gouvernance appliqués à la sélection, à la conception, au développement et au lancement de produits, services et infrastructures au sein du groupe France Telecom.

Chaque phase TTM commence et se termine par une décision go / no go formelle, appelée un **jalon**. Il existe six jalons, numérotés de T-1 à T4 (soit: T-1, T0, T1, T2, T3, T4 *cf. page suivante*). Les jalons sont matérialisés par des revues de jalon (par le biais de réunions, conférences téléphoniques/internet, et échanges de mails) et des décisions traçables, mentionnées dans un compte-rendu de revue.

a) les objectifs

TTM est le processus utilisé pour la sélection, la conception et le lancement des produits et infrastructures. TTM s'applique à tous les projets de produits et d'infrastructure aux niveaux du groupe, des pays et des Business Unit. TTM définit les règles pour la gestion de ces projets : jalons, livrables, organisation et gouvernance. TTM ne définit pas la manière dont les tâches sont exécutées dans les départements contributeurs.

TTM est conçu et mis en œuvre au service du business. Son objectif est d'optimiser le délai de mise sur le marché, la qualité et l'expérience client, la création de valeur, les dépenses d'investissement des projets de développement de produit et d'infrastructure. Il facilite par ailleurs le développement de produits et infrastructures multi-pays.

L'ILM (Suivi du cycle de vie du produit) étend TTM jusqu'à la fin de vie des produits. Roadmap Entry définit de quelle manière les concepts (futurs produits) rentrent dans la roadmap.

Cette méthode a permis de réduire le temps de lancement d'un produit de 41% depuis 2005.

b) les facteurs clés de succès de TTM

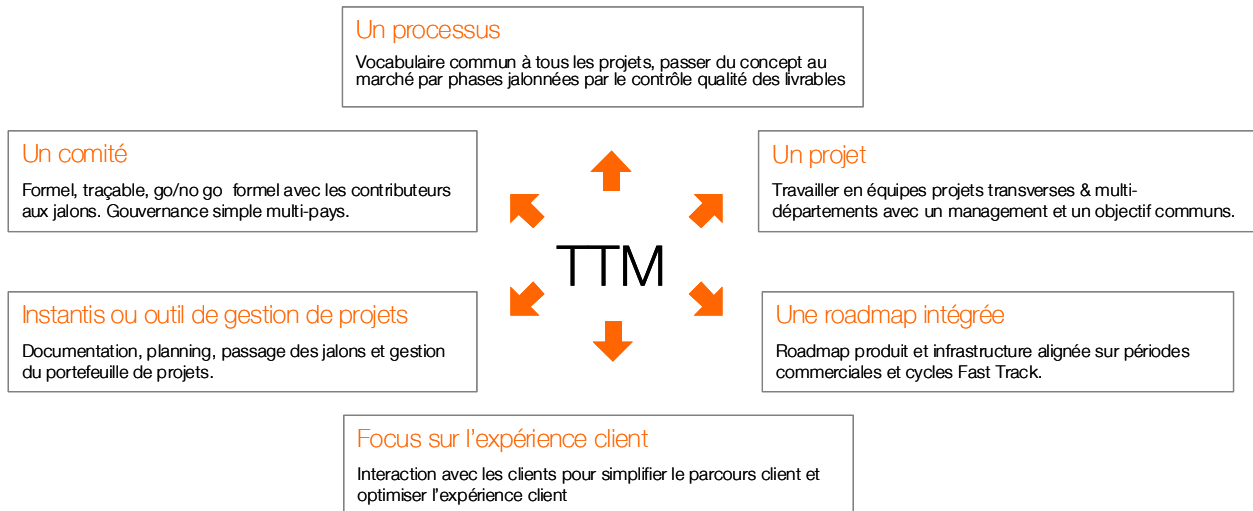


Illustration 16 : les six facteurs clés de succès

Les facteurs clés permettant de s'assurer la réussite d'un projet sont :

- travailler en mode projet avec une équipe projet encadrée par un chef de projet avec un objectif à tenir
- utiliser un processus simple avec :
 - o des projets phasés
 - o une check-list de livrables
 - o des contrôles qualité à chaque jalon
 - o des options TTM adaptées à la complexité des projets
- décider vite et bien en comité pour
 - o prendre les décisions aux jalons
 - o engager les ressources, OPEX📖 et CAPEX📖 de tous les contributeurs
 - o déléguer pour être efficace
- focus sur l'expérience client
 - o impliquer les clients de l'idée au produit
 - o focus sur la simplicité du parcours client et la qualité
 - o le champion client exécute le plan expérience client
- gérer la roadmap
 - o priorisée
 - o compatible avec les ressources
 - o mise à jour régulièrement en fonction du marché
- utiliser Instantis ou un autre outil pour
 - o partager et approuver les livrables
 - o gérer le planning et les ressources
 - o gérer le portefeuille de projets

c) les phases et jalons

Un projet TTM s'exécute en cinq phases. Chaque phase commence et se termine par une décision go/no go formelle, appelée **jalon**. Les six jalons, numérotés de T-1 à T4, se décomposent comme suit :

- T-1 ou revue de pré-opportunité : initialisation du projet avec l'affectation formelle de l'équipe projet en charge de l'étude d'opportunité.
- T0 ou revue d'opportunité : lancement du projet avec l'accord formel de la MOA sur le périmètre : cadre fonctionnel général, coûts, délais...
- T1 ou revue de conception : accord pour démarrer la phase de réalisation.
- T2 ou revue de développement : accord pour démarrer la phase de recette ou d'expérimentation.
- T3 ou revue de lancement sur le marché : accord pour ouverture généralisée du service, du produit ou de l'infrastructure ; transfert en activité récurrente ; passage de l'imputation financière de CAPEX en OPEX.
- T4 ou revue de bilan : fin du projet et validation du transfert de la responsabilité aux équipes des opérations.

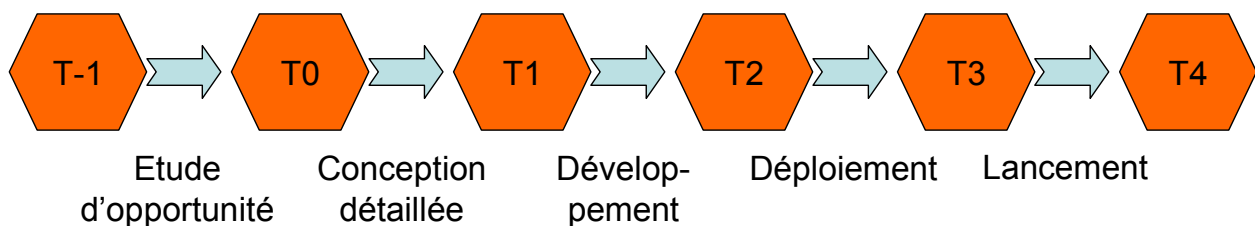


Illustration 17 : les jalons TTM

Le Comité TTM nomme l'équipe projet et prend, à chaque jalon, la décision de go/no go. Le comité comprend un représentant de tous les contributeurs au projet. Il valide formellement les OPEX, les CAPEX ainsi que les ressources humaines requises dans le plan projet.

Les jalons sont matérialisés par des revues de jalon (qui peuvent être des réunions, conférences téléphoniques/internet, ou échanges de mails) et des décisions traçables, mentionnées dans un compte-rendu de revue. Les check-lists définissent les livrables du projet à chaque jalon. Le chef de projet TTM est responsable de la qualité des livrables. Le Comité TTM garantit que la qualité est contrôlée pendant la préparation de la revue du jalon.

4.2. Les passages obligatoires du projet

Le passage des jalons TTM, marquant les étapes d'un projet chez France Telecom, sont sous la responsabilité du chef de projet, lequel demande ma contribution pour la rédaction des différents livrables. À titre personnel, je devrai passer, en tant que MOE📖, devant différentes commissions de validation dans le cadre de mon projet : CVAT📖, CI2A📖 et COGIT📖.

a) Le CVAT

Pour déployer, généraliser une nouvelle application ou une nouvelle version d'une application existante, il est nécessaire de faire **valider l'architecture technique** par le Comité de Validation des Architectures Techniques (CVAT📖). Le dossier d'architecture technique est un livrable attendu du processus TTM. Il conditionne le déclenchement du processus d'investissement. Le CVAT📖 est une organisation transverse à France Telecom qui fait appel aux compétences de l'ensemble des équipes d'architectes techniques du Groupe : le "collège des architectes techniques". Cette instance décisionnelle rend des avis applicables sur les architectures techniques qui lui sont présentées. Chaque grande entité (Sifac, Sires, SI📖 Client, DPS, Orange, etc.) est représentée au CVAT📖.

L'instruction se déroule en 3 phases :

- Le pré-CVAT📖 est l'instance technique qui effectue l'analyse des dossiers. Ce pré-comité regarde tout particulièrement la conformité de l'architecture technique présentée aux programmes Archimède et Platon📖. Un «porteur» est nommé pour chaque dossier, qui prépare une Proposition de Relevé de Décision (PRD). Celle-ci sert de trame à l'examen du dossier en pré-CVAT et est envoyée au chef de projet. Cette PRD sera ensuite présentée en CVAT📖. Les participants du pré-CVAT📖 sont des architectes techniques issus du collège des architectes, les urbanistes d'Information System Enterprise Architecture (ISEA), les MOE📖.
- Le CVAT📖 est présidé par le directeur de DDSI et le directeur de DOSI📖. Les participants sont des architectes issus du collège des architectes, les urbanistes d'ISEA. Les représentants de la maîtrise d'œuvre des applications présentées sont invités.
- Enfin, un Relevé de Décision, signé du directeur de DDSI et du directeur de DOSI📖, trace la validation de l'architecture technique du dossier présenté.

Le passage en pré-CVAT (PCVAT) puis en CVAT📖 se déroule devant un collège d'architectes devant lequel il faut démontrer que la solution retenue est **pérenne** et **s'intègre** bien à l'architecture actuelle du système d'informations du groupe [1].

Le passage en CVAT📖 et donc l'acceptation du DAT **conditionne la construction de l'infrastructure cible** de mon projet et le lancement des commandes. Comme nous l'avons vu au chapitre précédent, les jalons TTM conditionnent l'avancement du projet et sont donc sous la responsabilité de l'équipe projet ; en tant qu'architecte, le DAT est cependant sous ma responsabilité. J'axerai donc la suite de ce mémoire sur les éléments m'ayant permis de **rédiger le DAT et de passer le CVAT📖 avec succès**.

b) Le CI2A

Pour implanter du matériel dans un Datacenter📖, il faut présenter son projet en **Comité d'Implantation des Applications et des Architectures (CI2A📖)**, en présentant ses **besoins en termes d'occupation au sol (m²) et d'énergie (kW)**. Le CI2A📖 désigne l'instance d'hébergement, vérifie que les besoins peuvent être satisfaits et les réserve.

c) Le COGIT

Pour implanter un projet dans un IAS📖, il faut présenter son projet en **Cellule Opérationnelle de Gouvernance des Infrastructures Techniques (COGIT)**, en indiquant ses besoins en ingénierie réseau et en outils d'exploitation (supervision, sauvegarde, stockage, accès distant, messagerie SMTP...).

4.3. Organisation hiérarchique

Le positionnement global des acteurs du projet MIO Backup dans l'entreprise France Telecom (FT) est représenté ci-dessous :

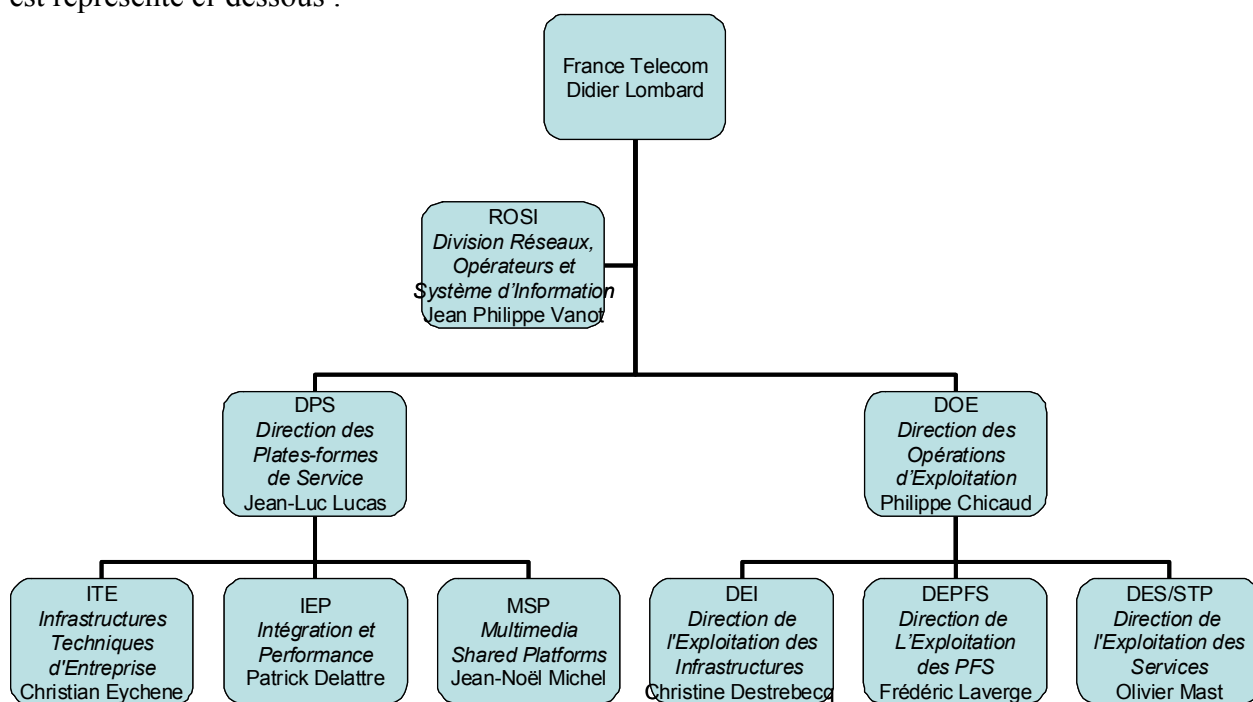


Illustration 18 : organigramme de la division ROSI

Comme nous l'avons vu au chapitre 1.5, la direction DPS est en charge de la conception des infrastructures et est à ce titre en charge du projet.

La direction DOE est plus particulièrement en charge de l'exploitation.

a) Services en charge de la gestion du projet

Ci-dessous une présentation d'une partie de l'organigramme de la direction DPS avec un focus (**en vert**) sur les principales entités et personnes avec lesquelles j'ai travaillé dans le cadre de ce projet :

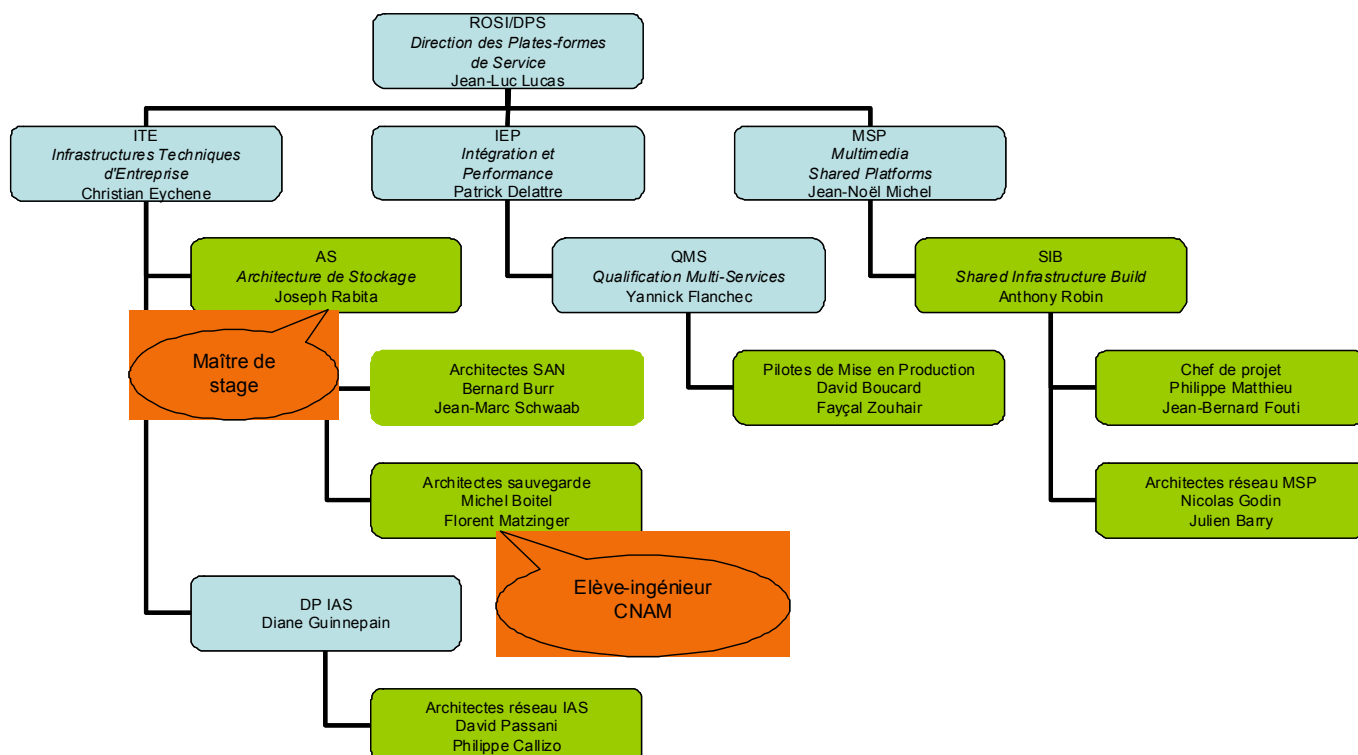


Illustration 19 : organigramme de la direction DPS

Ainsi, le chef de projet rattaché à FT/ROSI/DPS/MSP/SIB (FT / RO&SI / DPS / Multimedia Shared Plate-forme / Shared Infrastructure Build).

La Maîtrise d'OuvrAge (MOA📖) est représentée par FT/ROSI/DPS/EM-DIR (FT / RO&SI / DPS / État Major Direction DPS).

La **Maîtrise d'Œuvre (MOE📖)**, à laquelle mon maître de stage et moi-même appartenons, est rattaché à FT/ROSI/DPS/ITE/AS (FT / RO&SI / DPS / Infrastructures Techniques d'Entreprise / Architecture de Stockage).

b) Services en charge de l'exploitation

Ci-dessous une présentation d'une partie de l'organigramme de la direction DOE avec un focus (**en vert**) sur les principales entités et personnes avec lesquelles j'ai travaillé dans le cadre de ce projet :

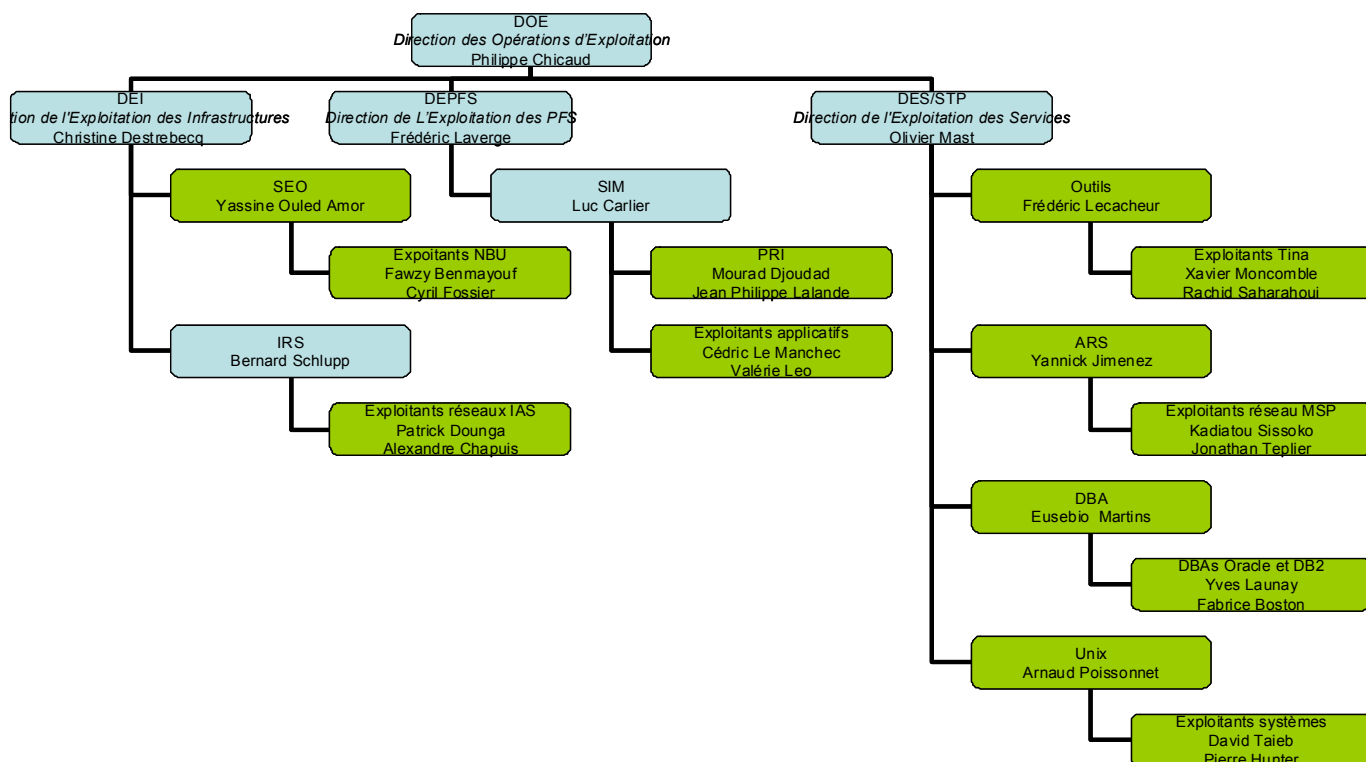


Illustration 20: organigramme de la direction DOE

On y retrouve en particulier les équipes d'exploitation suivantes :

DOE/DES/STP/OUTILS qui gère les outils (sauvegarde, supervision...) sur les serveurs de MDSP📖, en particulier l'infrastructure TiNa📖 et STK📖.

DOE/DES/STP/DBA📖 qui gère les bases de données Oracle📖 et DB2📖 de MDSP📖.







DOE/DES/STP/UNIX qui gère l'infrastructure SAN📖 Orange et les OS (systèmes d'exploitation) des serveurs MDSP📖.

DOE/DEPFS qui gère les applicatifs métiers sur les serveurs MDSP📖.

TMC/INFRA puis DEI/SEO qui gère l'infrastructure NBU📖 dans les IAS📖 (SAN📖 IAS, NBU et DXi📖).

4.4. Activités confiées

Mes activités sur ce projet sont les suivantes :

- **Réaliser un inventaire** du périmètre,
- **Proposer une infrastructure NBU**  **mutualisée** incluant la mise en œuvre d'une infrastructure de sauvegarde sur disques à coût réduit à base de VTL  ; je devrais décrire cette architecture dans un DAT à défendre en CVAT ,
- **Proposer une évolution de l'architecture du SAN**  **de Backup** en collaboration avec les architectes stockage,
- **Proposer une solution généralisée, optimisée et moins coûteuse** pour les sauvegardes de **base de données**,
- **Définir les procédures de migration** de TiNa  vers NBU ,
- **Commander le matériel** (serveurs, robotiques, périphériques...) **et les licences logicielles** pour l'infrastructure,
- **Piloter le déploiement** de l'infrastructure,
- **Assurer une bonne communication** avec tous les interlocuteurs afin de leur présenter les changements apportés et leurs impacts,
- **Assurer le support** sur tous les problèmes techniques rencontrés lors de la migration,
- **Assurer la réversibilité de la connaissance** pour assurer la maintenance récurrente.

5. ANALYSE DE L'EXISTANT

Selon le fournisseur Symantec [3], afin de concevoir une architecture de sauvegarde, il faut connaître :

- les serveurs à sauvegarder,
- la volumétrie à sauvegarder,
- le type de données (afin de déterminer son taux de compression ou de déduplication📖)
- la méthode de sauvegarde des bases de données,
- le nombre de fichiers à sauvegarder (pour l'estimation de la taille du catalogue),
- à quelle fréquence les sauvegardes seront exécutées
- les rétentions (combien de temps les sauvegardes sont conservées)
- la configuration réseau / la répartition géographique,
- les fenêtres de sauvegarde (les plages horaires pendant lesquelles on peut sauvegarder),
- le besoin d'externalisation
- l'évolution de tous les paramètres ci-dessus...

J'ai ainsi dû conduire différents **inventaires** afin de **déterminer le périmètre** précis du projet et de **concevoir la solution de sauvegarde** adaptée.

J'ai demandé aux équipes d'exploitation :

- de me fournir les extraits «tina_acct» des catalogues TiNa📖 (*voir un extrait dans l'annexe 3*),
- de lancer des requêtes d'interrogation sur les serveurs ACSLS📖 (*voir un exemple dans l'annexe 4*),
- de me fournir le nombre d'objets et d'instance enregistrés dans les catalogues TiNa,
- ...

J'ai demandé aux MOE📖 des différentes applications :

- de me fournir une estimation de la volumétrie actuelle et à venir,
- de renseigner les EB📖 de sauvegarde,
- de commander les licences NBU📖 nécessaires,
- ...

J'ai demandé aux architectes de présenter :

- les réseaux IP et SAN📖,
- le PRA📖 applicatif,
- les évolutions à venir de l'architecture de la plate-forme,
- ...

5.1. Audit du périmètre à sauvegarder

L'analyse des fichiers «tina_acct» et des résultats des requêtes ACSLS m'a permis d'établir cet état des lieux présentant le nombre de clients et l'infrastructure robotique actuellement en place :

Périmètre licences TiNa actuelles (réf. Décembre 2007)	Nombre
storage nodes	50
clients Windows	51
clients Linux	698
client Unix (Solaris et AIX)	207
bases Oracle	49
bases DB2 (service LDAP)	6

Tableau III : répartition des agents TiNa installés



Tableau IV : infrastructure actuelle



On me demande non seulement de **reprendre l'existant**, c'est-à-dire ce qui est sauvegardé sous TiNa, mais également de pouvoir **accueillir les nouveaux projets** sur une même infrastructure mutualisée. De plus, il faut distinguer le besoin immédiat du besoin à terme

pour la volumétrie des données. En effet, les bases de données, qui constituent l'essentiel de la volumétrie, ne sont que partiellement pleines. On peut les apparenter à une coquille initialement vide (l'enveloppe de la base de données) qui se remplit au fur et à mesure (l'occupation actuelle) jusqu'à occuper au maximum la taille de l'enveloppe, laquelle correspond à l'estimation de la volumétrie maximale qui pourrait être occupée en régime de croisière. Le besoin immédiat est donc constitué par l'occupation actuelle, mais le besoin à terme est constitué par l'enveloppe totale. **Dans un souci de pérennité, j'ai considéré le besoin à terme.**

Le besoin à terme global (c'est-à-dire de l'ensemble des bases de données sur lesquelles se concentre l'essentiel de la volumétrie utile et donc à sauvegarder) exprimé par les architectes applicatifs et retenu est de 5933 Go par environnement actif. Il y a donc $3 \times 5\,933 = 17\,799$ Go de données à sauvegarder sur la plate-forme MDSP (3 car les données sont présentes sur l'environnement de production actif (Aubervilliers ou Vélizy) et sur chaque pré-production).

Ces chiffres seront repris au chapitre 6.2 afin de dimensionner l'architecture.

5.2. L'environnement technique

Concernant l'environnement technique, les systèmes d'exploitation et applications suivants sont à prendre en compte, car actuellement présents au titre des clients ou seront présents en tant qu'infrastructure à déployer :

	Systèmes d'exploitation	Applications
Master Serveurs	❑ Linux Red Hat AS,	❑ NetBackup MASTER SERVEUR.
Media Serveurs	❑ Linux Red Hat AS, ❑ Solaris.	❑ NetBackup MEDIA SERVEUR.
Clients	❑ AIX, ❑ Solaris, ❑ Linux Red Hat AS, ❑ Windows 2000 et 2003, architecture 32 et 64bits.	❑ NetBackup SAN MEDIA SERVEUR, ❑ NetBackup CLIENT, ❑ Oracle 9i et 10G, ❑ DB2 8.1, ❑ MySQL 5.

Tableau V : environnement technique

J'ai vérifié que le produit NBU était bien **supporté par l'éditeur Symantec**, en consultant la matrice de compatibilité[4] et les release notes des versions 6.x, et **qualifié par mon équipe (la MOE sauvegarde)** pour tous les clients présents sur la plate-forme. J'ai fourni aux exploitants en charge du déploiement les packages Platon à déployer sur les clients :

Version Platon de l'OS	Version OS	NBU 6 MP4 Client	NBU 6 MP4 Media
AIX G4R0	AIX 5.2 ML2	TRF10010	TRF10052
AIX G4R1	AIX 5.2 ML4	TRF10010	TRF10052
AIX G5R1	AIX 5.3 ML3	TRF10010	TRF10052
Linux G4R0	AS 2.1	TRF9940	TRF10029
Linux G6R0	AS 3	TRF9940	TRF10029
Linux G7R0	AS 4	TRF9940	TRF10029
Solaris G5	Solaris 9	TRF10083	TRF10041
Windows G3R3	Windows 2000 SP4	TRF9690	TRF9792
Windows G4R1	Windows 2003 SP1 SP2	TRF9690	TRF9792

Tableau VI : packages Platon à déployer

J'ai également transmis le pré-requis, relevé lors de la qualification de NetBackup 6, concernant la mise à jour de la librairie libstdc sur les versions Platon de RedHat :

- installer le package compat-libstdc++-6.2-2.9.0.16.i386.rpm sur les serveurs Linux Platon G4 (RedHat AS 2.1),
- installer le package compat-libstdc++-devel-7.3-2.96.128.i386.rpm sur les serveurs Linux Platon G6 et G7 (RedHat AS 3 et 4).

5.3. Le découpage en environnement et DMZ

Chaque service du projet MDSP est hébergé par des serveurs regroupés par **environnement** :

- INT (Intégration)
- PPHOM (Pré-production homologation)
- PPMCE (Pré-production maintenance)
- PPERF (Pré-production performance). *Ce dernier environnement sera décommissionné en cours de programme MIO ; je n'en tiendrai donc pas compte dans l'étude.*
- PROD (Production). *Ce dernier environnement est réparti sur les deux sites, les bases de données étant répliquées via SRDF du site actif (Aubervilliers par défaut) vers le site de PRA (Vélizy par défaut).*

Chaque environnement est lui-même découpé en zones ou **DMZ** :

- une zone d'**administration**, qui est une zone transverse dédiée à l'administration technique de l'application. Cette zone héberge tous les composants nécessaires à l'administration (supervision, centralisation des logs, bastions d'administration ...). Cette zone est également utilisée pour administrer tous les composants de l'infrastructure.
- une zone de relais dite «**Publique**» en frontal de l'Internet et du Réseau d'Accès Partenaire, avec un double objectif :
 - o un objectif de relais des flux vers la DMZ Privée,
 - o un objectif de sécurité : filtrage des flux depuis Internet vers la zone de traitement.
- une zone de traitement dite «**Privée**» où l'on retrouve l'application proprement dite, qui constitue le cœur de l'architecture.
- une zone de données dite «**Secure**» ou «**Trusted**» où sont localisées les données de la plateforme. Cette zone Trusted accède en direct au RSC.
- une zone de relais interne, dite «**Publique Interne**» ou «**Publique RSC**», en frontal du RSC dont l'objectif est similaire à la zone de relais évoquée ci-dessus.

D'après ce principe, chaque serveur (hors zone d'administration) a donc au moins une interface sur le réseau de production et une interface sur le réseau d'administration, et communique avec la DMZ d'administration à travers son interface d'administration uniquement.

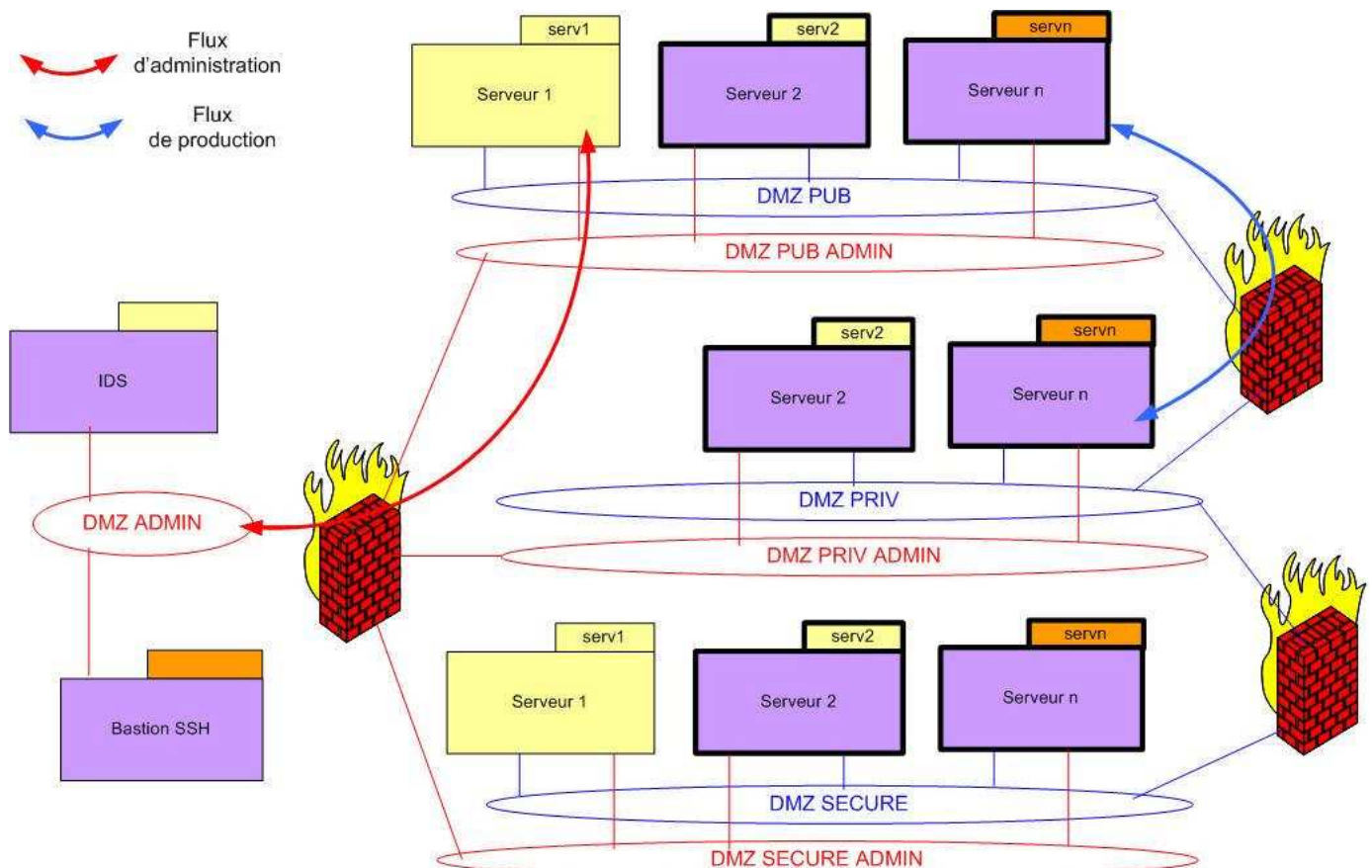


Illustration 21 : schéma logique détaillant les flux de production et d'administration

5.4. L'intégration dans l'environnement cible : les IAS

Parmi les projets du programme MIO, le plus impactant, car structurant reste certainement le projet de migration réseau pour intégrer les IAS, afin de mutualiser les ressources et en particulier le service de sauvegarde. Dans le cadre de mon projet, je dois proposer une architecture de sauvegarde mutualisée entre MDSP et les IAS, et donc analyser l'infrastructure de sauvegarde dans les IAS. Ce travail sera trivial pour moi, car j'avais participé à la mise en place de la première architecture de sauvegarde dans les premiers IAS dont Aubervilliers.

Voir l'annexe 1 pour plus d'informations sur les IAS et les zones qui la composent.

Ci-dessous l'infrastructure de sauvegarde d'un IAS :

- le master NBU, la ou les robotique(s), le switch SAN sont implantés dans la ZSV,
- l'accès aux serveurs se fait via un VLAN nommé «administration sauvegarde» au travers de bastions sécurisés nommés TDIMG (solution d'ICA),
- les flux de sauvegarde transitent via un VLAN nommé «service sauvegarde» lequel dessert toutes les DMZ où sont implantés les clients à sauvegarder.

Les autres serveurs offrant des services tels que DNS, NTP, SMTP, supervision, rebond... que j'utiliserai dans le cadre du déploiement ou de l'exploitation de l'infrastructure de sauvegarde, sont implantés dans la zone d'exploitation.

Ci-dessous une représentation de l'intégration réseau MDSP-IAS prévue par les architectes réseau à ce stade du projet :

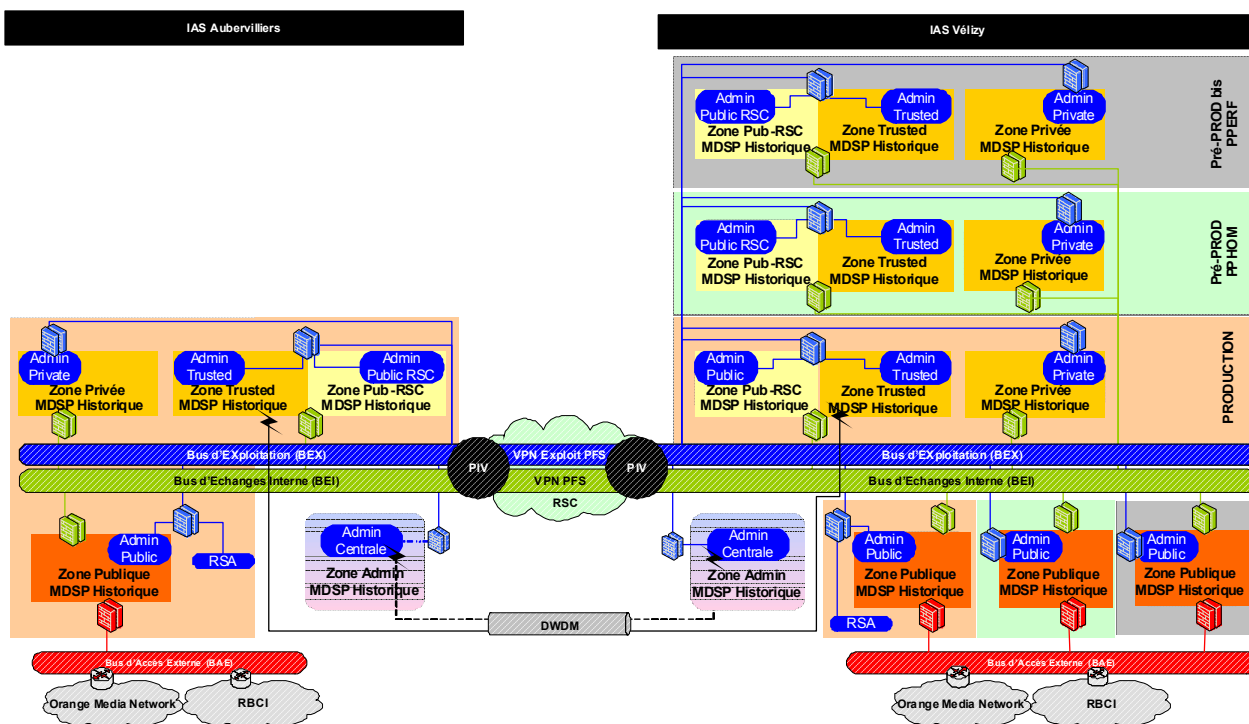


Illustration 22 : schéma prévisionnel d'intégration des applications MDSP dans les IAS

5.5. Audit du SAN de backup

Le SAN de backup de la plate-forme de production MDSP est en réalité un SAN de backup mutualisé appelé «SAN Backup Orange». Il englobe le périmètre MDSP, mais également d'autres projets plus ou moins conséquents et sensibles tel que «Content Billing».

Cette architecture doit être remise en cause car :

- Ce SAN est composé de switches Brocade 12000 et 24000, lesquels ne seront plus supportés par Brocade en 2010,
- les exploitants SAN Orange et IAS appartiennent à des équipes distinctes.

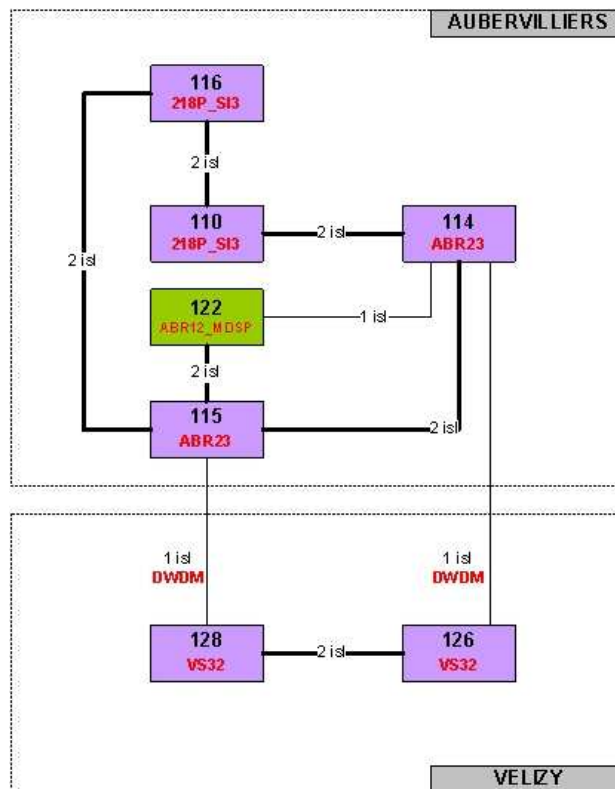


Illustration 23 : fabric étendue du SAN de backup en production

Le SAN de backup de la plate-forme de pré-production MDSP est en réalité un SAN mutualisé entre sauvegarde et stockage. Aucune évolution de celui-ci n'est prévue.

La fabric de backup IAS sur le site d'Aubervilliers est actuellement composée d'un seul switch Brocade 4100 (32 ports).

Il n'y a actuellement pas de SAN de backup IAS sur le site de Vélizy ; celui-ci sera donc **à construire**.

On configure le zoning par port sur le SAN Backup Orange, et par WWN sur le SAN Backup IAS.

5.6. Contraintes et risques du projet

Contraintes	Oui	Désignation
Date délais impactant d'autres projets (applis, infra...)	■	Disponibilité des salles ABS31 et VS33, disponibilité de l'IAS de Vélizy, interconnexion réseaux IP et SAN entre MDSP et IAS (sur les 2 sites : Aubervilliers et Vélizy).
Contraintes d'interface	<input type="checkbox"/>	
Date de T2 incompressible	■	30 Septembre 2008 – L'infrastructure doit être opérationnelle pour accueillir les nouvelles applications ainsi que les plates-formes migrées.
Date de T3 incompressible	■	09 Mars 2009 – Ouverture du service.
Contraintes de Budget	<input type="checkbox"/>	
Contraintes de moyens disponibles (matériel infra...)	■	Lien réseau inter-sites pour réplication des cartouches entre les VTL des deux sites.
Contraintes matérielles	■	Disponibilité de la GA pour les DXi 7500 (prévu le 27 Mai 2008 ; action Quantum).
Contraintes exprimées dans les travaux en amont	<input type="checkbox"/>	
Contraintes de Production des contrats d'interface	<input type="checkbox"/>	
Contraintes de déploiement	<input type="checkbox"/>	
Contraintes de MPP	<input type="checkbox"/>	
Contraintes de qualification	■	Agent DB2 à qualifier pour NBU.
Contraintes de MEP	■	Le service de sauvegarde proposé par la nouvelle architecture doit être disponible au T2 (30 Septembre 2008).
Autres, à compléter		

Tableau VII : contraintes exprimées dans le PMP

Pour ce projet, nous avons identifié les types de **risques** suivants :

- **Planning** : forte adhérence avec d'autres projets, non-respect des jalons compte tenu du nombre d'intervenants sollicités, retard dans la disponibilité de l'IAS [livre] de Vélizy, retard de livraison du matériel (en particulier la GA du DXi [livre] 7500), retard dans la disponibilité des salles ABS31 et VS33, retard dans la bascule réseau dans l'IAS [livre] (projet de migration réseau du programme MIO), retard dans la rédaction des EB [livre] par les MOE [livre] applicatives...
- **Technique** : disponibilité des liens réseau inter-sites, impossibilité d'interconnecter les SAN [livre] de sauvegarde, manque d'espace disque sur les serveurs, exploitation du serveur Tina lorsque celui-ci a basculé dans l'IAS [livre], en particulier pour la prise en charge des restaurations...
- **Stratégique** : performances attendues de la solution non atteints (taux de déduplication [livre], débit...)
- **Organisationnel** : multiplicité et disponibilité des intervenants, coordonner les 2 équipes d'exploitants outils (STP et TMC), vaincre les réticences des exploitants au changement...

Les risques de décalage de planning ont été en toute logique suivis par le chef de projet, mais m'ont néanmoins directement impacté. Le chef de projet et moi-même avons dû gérer les problèmes organisationnels pour faire réaliser les actions de migration aux différents exploitants. De par ma fonction d'architecte technique, j'ai été plus particulièrement **confronté aux risques techniques et stratégiques** dont j'avais **anticipé** les parades ci-dessous :

- **disponibilité des liens réseau inter-sites** : Le Metropolitan Area Network a été livré après la fin du projet. J'ai mis en place avec les architectes réseaux une réplication via les liens DWDM [livre], déjà utilisés pour le SAN [livre] de Backup Orange.
- **disponibilité de la General Availability (GA) du DXi [livre] 7500** : Mise en place d'un «plan B» pour permettre la sauvegarde de l'IAS [livre] de Vélizy sur un DXi [livre] 5500 sur Aubervilliers via les liens DWDM [livre].
- **impossibilité d'interconnecter les SAN [livre] de sauvegarde** : voir le chapitre 5.4
- **manque d'espace disque** sur les serveurs : j'ai commandé un audit sur les serveurs afin de permettre au plus tôt de libérer et/ou d'allouer de l'espace au besoin
- **exploitation du serveur Tina** lorsque celui-ci a basculé dans l'IAS [livre], en particulier pour la prise en charge des restaurations : j'ai veillé auprès des architectes réseau que l'équipe STP en charge de l'infrastructure Tina pourra toujours l'exploiter et assurer les restaurations. Ceci a évité de former les exploitants TMC à l'outil Tina qu'ils n'auraient exploité que quelques mois.
- **performances** de la solution non atteintes : un capacity planning a été mis au point afin de superviser les DXi [livre] et d'anticiper les éventuelles commandes d'extension. J'ai dû en tenir compte dans mes calculs de charge, par exemple en introduisant un taux de déduplication [livre] plus faible que celui donné par le fournisseur.

J'ai pour ma part émis les exigences suivantes auprès des autres acteurs du projet :

- **accès** : j'ai veillé à ce que les architectes réseaux garantissent le fait que les exploitants doivent toujours pouvoir accéder facilement aux serveurs dont ils ont la charge, en particulier les exploitants TiNa📖 doivent pouvoir continuer à exploiter cet outil pendant et après la bascule réseau dans l'IAS📖 (*cf. chapitre 6.4*),
- **licence** : tout serveur sauvegardé via TiNa📖 ou NBU📖 doit disposer de la licence nécessaire (*cf. chapitre 7.2*),
- **EB📖** : toute demande de mise en sauvegarde doit être tracée dans un document intitulé «Expression de Besoin» de sauvegarde détaillant les exigences de la MOE applicative quant à ses données (*cf. chapitre 5*),
- **RMANOO** : afin de se mettre en conformité avec les préconisations groupe, j'ai demandé aux DBA📖 de migrer les packages de scripts RMAN de RMANOO vers RMAN Lite (*cf. chapitre 7.7*).

6. DÉFINITION DE LA SOLUTION TECHNIQUE

6.1. Architecture générale

Les architectures TiNa et NBU différant quelque peu dans leur architecture et dénomination, j'ai **largement communiqué** auprès des différents interlocuteurs du projet afin de présenter les différences de fonctionnement et de veiller à ce que tous utilisent un vocabulaire commun.

Les clients TiNa envoient les données à sauvegarder et les méta-données à un serveur TiNa via IP. Ils seront migrés en clients IP NBU, lesquels envoient les données à sauvegarder à un media serveur NBU (on ajoutera le qualificatif «d'infrastructure» afin de mieux les distinguer des SAN medias serveurs, car ils font partie de l'infrastructure de sauvegarde NBU) et les méta-données à un master serveur NBU via IP.

Les storage nodes Tina envoient les données à sauvegarder à la robotique via le SAN et les méta-données à un serveur TiNa via IP. Ils seront migrés en SAN medias serveurs NBU, lesquels envoient les données à sauvegarder à la robotique via le SAN et les méta-données à un master serveur NBU via IP.

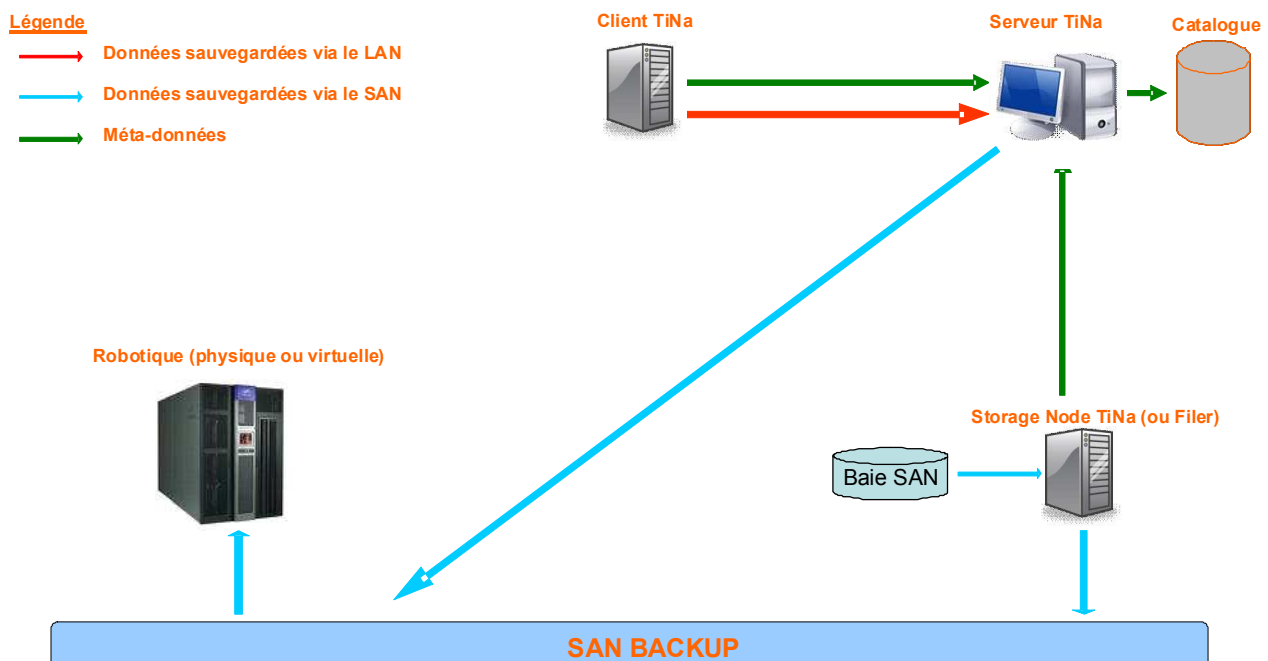


Illustration 24 : schéma fonctionnel d'une sauvegarde TiNa

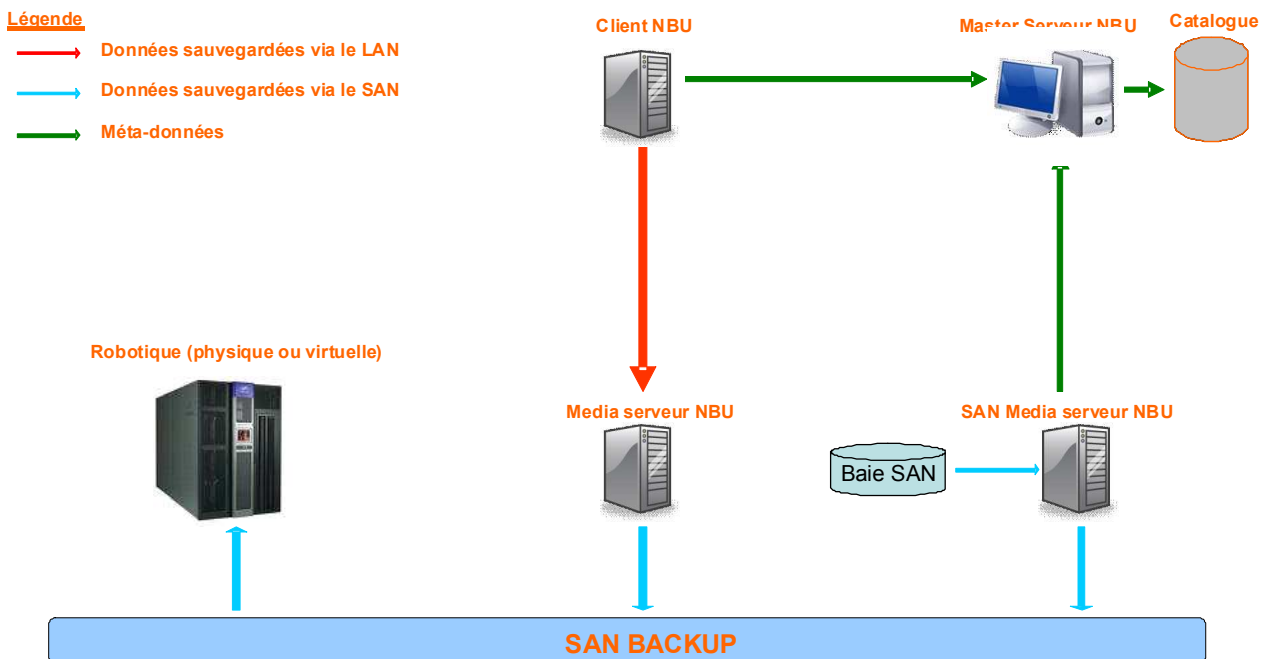


Illustration 25 : schéma fonctionnel d'une sauvegarde NBU

Tous les serveurs sauvegardés avec l'outil NBU envoient donc leurs méta-données à un master NBU qui héberge le catalogue.

Le but consiste à remplacer les serveurs TiNa par un(des) master(s) et des medias serveurs NBU.

La plate-forme MDSP étant décomposée en plusieurs DMZ, l'architecture standard de sauvegarde dans les IAS (media serveur dans la ZSV) ne peut convenir. En effet, les firewalls d'accès aux différentes DMZ et surtout le firewall d'accès à la ZSV constitueraient le goulet d'étranglement de la solution.

J'ai donc proposé en CVAT une autre architecture, laquelle permet de déporter les medias serveurs dans les DMZ. Comme présenté au chapitre 4.2, le Comité de Validation des Architectures Techniques constitue le passage obligatoire pour toute nouvelle implantation d'architecture afin de s'assurer qu'elle respecte les préconisations groupe et donc sa bonne intégration dans l'architecture déjà présente [1].

Après concertation avec les architectes réseaux, ceux-ci ont exprimé les contraintes suivantes sur les flux :

- pas de communication directement entre les DMZ publique et trusted (passage obligatoire par la DMZ privée pour des questions de sécurité),
- aucun filtrage entre les DMZ trusted et publique RSC,
- aucun filtrage entre les environnements PPHOM et INT,
- aucun filtrage entre les environnements PPMCE et PPERF,
- préférence pour que les serveurs de la DMZ d'administration communiquent avec la DMZ privée ou publique,
- pas de communication entre les différents environnements ; seule la DMZ d'administration est commune à tous les environnements.

Dans un souci de simplification, on peut considérer, aux sens réseau et sauvegarde (mais non applicatif), que :

- les environnements PPHOM et INT ne forment qu'un seul ensemble,
- les DMZ📖 trusted et publique RSC📖 ne forment qu'un seul ensemble.

Je considérerai donc dans la suite du document qu'il y a :

- 4 environnements : PPHOM-INT, PPMCE-PPERF et les 2 PROD,
- 3 DMZ📖 par environnement : privée, publique et secure- (ou trusted-) publique RSC.

J'ai finalement proposé :

- par site : le maintien (sur Aubervilliers) ou le déploiement (sur Vélizy) d'un master📖 afin d'avoir un **catalogue unique** par site.
- par DMZ📖 : le déploiement d'un media serveur📖 (d'infrastructure) afin que les données sauvegardées ne traversent pas de firewall📖.
- par environnement : la création d'une VTL📖 pilotée par le media serveur📖 de la DMZ📖 privée.
- pour l'IAS📖 de Vélizy (hors MDSP📖) : un media serveur📖 avec sa VTL📖 dédiée.

Cette solution permet de limiter les données devant traverser un(des) firewall(s)📖 aux flux :

- de méta-données📖,
- de mise à jour de la base EMM et de commande robotique (réservation de lecteurs, demande de montage et de démontage de bande).

Des Demandes d'Ouvertures de Flux (DOF) devront être émises pour les besoins de flux suivants :

- tous les serveurs doivent communiquer avec le master📖,
- tous les medias serveurs doivent communiquer avec le media serveur📖 de la zone privée qui pilote la VTL📖,
- les serveurs de la zone d'administration communiquent avec le media de la DMZ📖 Privée (c'est-à-dire le moins chargé mais proposant du disk staging📖), la quantité de données à sauvegarder dans cette DMZ📖 ne justifiant pas d'ajouter de media serveur📖 dédié.

Les flux utilisés sont :

- TCP 1556 (pbx_exchange) entre le serveur hébergeant la base EMM (le master📖 serveur) et les medias serveurs, et entre les medias serveurs des DMZ📖 Publique et Trusted avec le media serveur📖 de la DMZ📖 Privée.
- TCP 13724 (vnetd) entre le master📖 et les clients, entre le media et ses clients, entre les medias serveurs des DMZ📖 Publique et Trusted avec le media serveur📖 de la DMZ📖 Privée.

Nous vérifierons dans le chapitre suivant que l'infrastructure proposée supportera la charge.

La répartition des clients par media serveur et DXi serait ainsi la suivante :

Site	Environnement	DMZ desservies	Nb de clients IP		Nb de SAN média		Média serveur
Vélizy	PPHOM+INT	Public	77	266	0	0	mbspvq347
Vélizy	PPHOM+INT	Admin et Private	95		0		mbspvq636
Vélizy	PPHOM+INT	Trusted et Pub RSC	38		0		mbspvq760
Vélizy	PPMCE+PPERF	Public	71	187	0	7	mbspvr364
Vélizy	PPMCE+PPERF	Admin et Private	89		0		mbspvr636
Vélizy	PPMCE+PPERF	Trusted et Pub RSC	27		7		mbspvr779
Vélizy	PROD	Public	128	298	0	13	mbspvp389
Vélizy	PROD	Admin et Private	141		0		mbspvp670
Vélizy	PROD	Trusted et Pub RSC	29		13		mbspvp787
Aubervilliers	PROD	Public	114	278	0	13	mbspap389
Aubervilliers	PROD	Admin et Private	136		0		mbspap666
Aubervilliers	PROD	Trusted et Pub RSC	28		13		mbspap785

Tableau VIII : répartition des clients par media serveur

6.2. Calculs capacitifs

Les rétentions en vigueur dans les IAS, et qui s'imposent donc à la plate-forme MDSP, sont les suivantes :

- 10 semaines pour le système et les applications, à raison d'une totale toutes les 4 semaines et des incrémentales les autres semaines,
- 16 jours pour les données, à raison d'une totale chaque week-end et des incrémentales les autres jours.



Les plages horaires autorisées pour les sauvegardes sont les suivantes :

- totales mensuelles : une semaine sur quatre, le samedi de 0h à 8h ou le dimanche de 0h à 8h,
- incrémentales hebdomadaires : le samedi de 0h à 8h ou le dimanche de 0h à 8h, sauf la semaine de la totale,
- totales hebdomadaires : le samedi de 0h à 8h ou le dimanche de 0h à 8h,
- incrémentales quotidiennes : tous les jours (sauf le jour de la totale) de 0 à 8h.





On aura donc en ligne :


- toujours 2 totales et 8 incrémentales en rétention 10 semaines,
- au maximum 3 totales et 13 incrémentales en rétention 16 jours.


Posons les variables suivantes par environnement :

- sc la volumétrie du système et des applications à sauvegarder sur les clients (rétention 10S),
- dc la volumétrie des données à sauvegarder sur les clients (rétention 16J),
- sm la volumétrie du système et des applications à sauvegarder sur les SAN  medias (rétention 10S),
- dm la volumétrie des données à sauvegarder sur les SAN  medias (rétention 16J).

Afin de dimensionner l'architecture SAN , on admettra pour les calculs suivants que :

- 15% des données varient chaque jour
- taux moyen de déduplication  = 7 (*le taux de déduplication dépend fondamentalement du modèle de données ; comme expliqué au chapitre 5.6, j'ai volontairement choisi l'hypothèse basse des taux de déduplication  observés sur les DXi 5500 déjà en production chez France Telecom*)
- nbclients = 265 pour chacun des 4 environnements
- nbsanmedias = 15 pour chacun des 4 environnements
- sc = nbclients * 5 Go = 1325 Go pour chacun des 4 environnements
- dc = nbclients * 3 Go = 795 Go pour chacun des 4 environnements
- sm = nbsanmedias * 10 Go = 150 Go pour chacun des 4 environnements
- dm = 5933 Go pour l'environnement de production actif (Aubervilliers **ou** Vélizy) et chaque pré-production.
- *un SAN media serveur  héberge une base de données et réciproquement une base de données est hébergée sur un SAN media serveur  (sauf la base OTAP DMC sauvegardée via IP au vu de sa faible volumétrie, mais je ne tiendrai pas compte de cette exception afin de simplifier les calculs qui suivent).*

On distinguera le cas nominal, où la production est active sur Aubervilliers (et donc les bases sont montées sur Aubervilliers, donc dm = 0 sur l'environnement de production de Vélizy), du PRA , où la production a basculé sur Vélizy (et donc les bases sont montées sur Vélizy).

Volumétrie qui sera stockée sur la VTL  d'Aubervilliers :

$$\begin{aligned} & 2*(sc+sm) + 8*(sc+sm)*0,15 + 3*(dc+dm) + 13*(sd+dm)*0,15 \\ & = 3,2*(sc+sm) + 4,95*(dc+dm) \\ & = 38\,023,6 \text{ Go} \\ & = 5,3 \text{ To déduplicué} \end{aligned}$$

Volumétrie qui sera stockée sur la VTL de Vélizy **en cas de PRA** :

$$\begin{aligned} & 3*(3,2*(sc+sm) + 4,95*(dc+dm)) \\ & = 114\,070,8 \text{ Go} \\ & = 15,9 \text{ To déduplicué} \end{aligned}$$

Volumétrie qui sera stockée sur la VTL de Vélizy **en fonctionnement nominal (hors PRA)** :

$$\begin{aligned} & 3*3,2*(sc+sm) + 4,95*(3*dc+2*dm) \\ & = 84\,702,45 \text{ Go} \\ & = 11,8 \text{ To déduplicué} \end{aligned}$$

Après réplication, volumétrie en ligne sur chaque site :

$$\begin{aligned} & = 38\,023,6 + 84\,702,45 \\ & = 122\,726,05 \text{ Go} \\ & = 17,12 \text{ To déduplicué} \end{aligned}$$

Calcul du débit nécessaire pour les sauvegardes totales (si elles s'exécutent toutes le même week-end) :

Sur Aubervilliers : $27\,945,6 / 16 = 1,7 \text{ To/h}$

Sur Vélizy : $84\,702,45 / 16 = 5,2 \text{ To/h}$

Les DXi ne pouvant absorber de tels flux (débit d'entrée de 800Go/h pour la version 5500, 1 To/h pour la version 7500), il faut donc trouver un moyen de **répartir les sauvegardes dans le temps**.

Une première répartition horaire (*représentée dans le tableau ci-dessous*) afin d'optimiser l'utilisation du DXi consiste à répartir les sauvegardes totales comme suit :

- 1/8 des sauvegardes du système et des applications le samedi de la 1^{ère} semaine, 1/8 le dimanche de la 1^{ère} semaine, 1/8 le samedi de la 2^{ème} semaine et ainsi de suite jusqu'à la 4^{ème} semaine
- la moitié des sauvegardes des données le samedi et l'autre moitié le dimanche ;

les 7/8 des sauvegardes incrémentales du système et des applications continuant à s'exécuter le week-end et les incrémentales des données les autres jours de la semaine.

Cette répartition des sauvegardes sera affinée dans le cadre du projet «MDSP backup infrastructure service management and evolution» que je piloterai après le T4 du projet, afin de gérer les nouveaux projets et la maintenance récurrente, et ne fait donc pas partie de ce mémoire.

[illegible]

Tableau IX : répartition temporelle des sauvegardes

Débit nécessaire pour les clients sur Aubervilliers :

- le samedi de 0h à 8h ou le dimanche de 0h à 8h :

$$(sc/8 + 7/8*sc*0,15 + dc/2 + dc/2*0,15) / 8$$

$$= (0,25625 * sc + 0,575 * dc) / 8$$

$$= 100 \text{ Go/h}$$

- tous les jours (sauf le jour de la totale) de 0 à 8h : $(dc*0,15) / 8 = 15 \text{ Go/h}$

Débit nécessaire pour les SAN medias serveurs sur Aubervilliers :

- le samedi de 0h à 8h ou le dimanche de 0h à 8h :

$$\begin{aligned} & (sm/8 + 7/8*sm*0,15 + dm/2 + dm/2*0,15) / 8 \\ & = (0,25625*sm + 0,575*dm) / 8 \\ & = 431 \text{ Go/h} \end{aligned}$$

- tous les jours (sauf le jour de la totale) de 0 à 8h : $(dm*0,15) / 8 = 111 \text{ Go/h}$

Débit nécessaire pour les clients sur Vélizy :

- le samedi de 0h à 8h ou le dimanche de 0h à 8h :

$$\begin{aligned} & 3 * (sc/8 + 7/8*sc*0,15 + dc/2 + dc/2*0,15) / 8 \\ & = 3*(0,25625*sc + 0,575*dc) / 8 \\ & = 299 \text{ Go/h} \end{aligned}$$

- tous les jours (sauf le jour de la totale) de 0 à 8h : $3*(dc*0,15) / 8 = 75 \text{ Go/h}$

Débit nécessaire pour les SAN medias serveurs sur Vélizy **en cas de PRA** :

- le samedi de 0h à 8h ou le dimanche de 0h à 8h :

$$\begin{aligned} & 3*(0,25625*sm + 0,575*dm) / 8 \\ & = 1293 \text{ Go/h} \end{aligned}$$

- tous les jours (sauf le jour de la totale) de 0 à 8h : $3*(dm*0,15) / 8 = 333 \text{ Go/h}$

Débit nécessaire pour les SAN medias serveurs sur Vélizy **en fonctionnement nominal (hors PRA)** :

- le samedi de 0h à 8h ou le dimanche de 0h à 8h :

$$\begin{aligned} & (3*0,25625*sm + 2*0,575*dm) / 8 \\ & = 867 \text{ Go/h} \end{aligned}$$

- tous les jours (sauf le jour de la totale) de 0 à 8h : $2*(dm*0,15) / 8 = 222 \text{ Go/h}$

Il apparaît donc que :

- en cas de PRA, le DXi de Vélizy ne pourrait absorber les flux de sauvegarde simultanés de la production et de la pré-production, il faudra donc désactiver les sauvegardes d'une pré-production. Ce passage en **mode dégradé** est validé par les applications car compatible avec leur PRA (en cas de crash du site d'Aubervilliers, la production bascule sur le site de Vélizy ; un des environnements de pré-production sera déménagé sur le nouveau site de PRA qui sera aménagé).

- en fonctionnement nominal, on ne peut exécuter en même temps les sauvegardes des SAN medias serveurs et des clients. La solution sera apportée par la fonctionnalité NBU de **disk staging** : celle-ci permettra ainsi d'utiliser en journée les baies de disque «tampon» des medias serveurs d'infrastructure pour les sauvegardes des clients IP, afin de «réserver» le DXi (et le SAN) pour les sauvegardes des SAN medias serveurs la nuit.

Le disk staging apporte les avantages suivants :

- **iso-fonctionnement** par rapport à l'infra TiNa (macro-multiplexage des clients IP),
- sauvegarde des clients IP sur le disk staging la nuit, puis **destaging** sur le DXi en journée ; ceci afin de ne pas interférer avec la sauvegarde des SAN medias serveurs sur le DXi la nuit,
- écriture de plusieurs jobs de clients IP en parallèle **sans activer le multiplexage** (qui a un effet négatif sur la déduplication),
- **tolérance de pannes** en cas d'indisponibilité temporaire du DXi ou du SAN.

6.3. Choix des éléments de l'infrastructure de sauvegarde

a) robotique

Les robotiques actuellement top-sourcées chez France Telecom sont :

- Adic/Quantum Scalar 100 et i2000
- STK SL 8500
- Quantum DXi 5500
- Fujitsu-Siemens CentricStor

J'ai tout d'abord comparé ces différentes solutions robotiques en termes de fonctionnalités :

	Traditionnelle	Virtuelle	
	Robots STK	Fujitsu Siemens CentricStor	Quantum Dxi
Compression	✓	✓	✓
Virtualisation		✓	✓
Déduplication			✓

Tableau X : comparatif des robotiques

J'ai ensuite proposé les différents scénarios suivants de consolidation et de mutualisation avec l'existant :

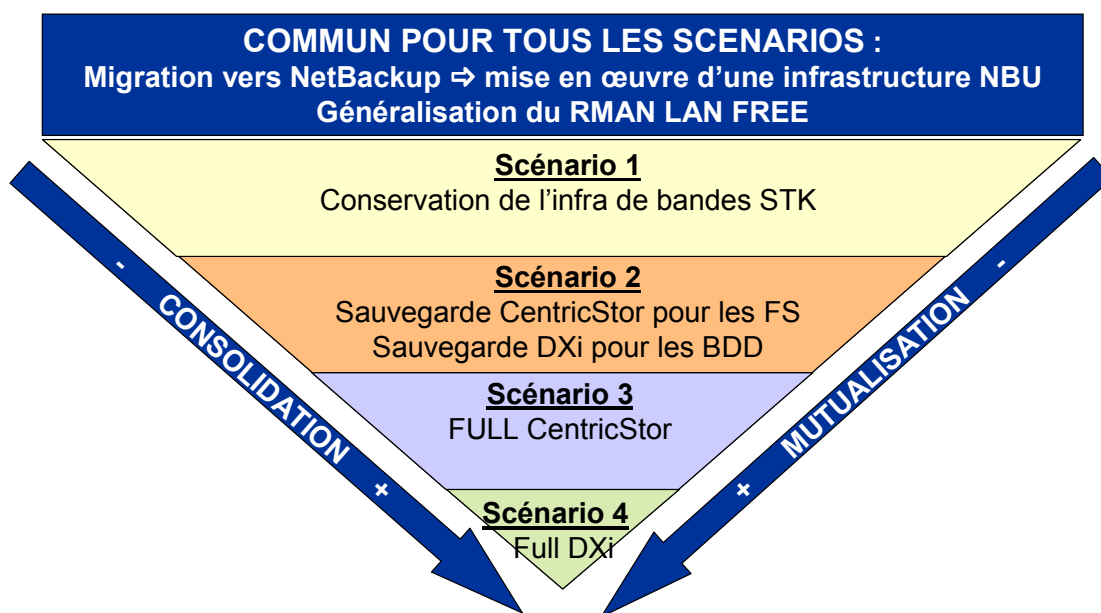


Illustration 26 : scénarios robotiques

Le scénario 1 de conservation des bibliothèques existantes présente 4 inconvénients :

- Nécessite l'acquisition de nouveaux lecteurs
- Non mutualisable avec l'infrastructure IAS (en VTL)
- Sauvegarde croisée (dite «en Y») via des liens DWDM
- Performance de compression de 2 :1

Le scénario 2 de sauvegarder les file systems des différents clients sur CentricStor et les bases de données sur DXi présente 3 inconvénients :

- Nécessite l'upgrade de l'infrastructure CentricStor
- Mutualisé sur les moyens DOSI
- Solution non homogène

Le scénario 3 de sauvegarder toutes les données sur CentricStor présente 5 inconvénients :

- Nécessite l'upgrade de l'infrastructure CentricStor (idem scénario 2)
- Mutualisé sur les moyens DOSI (idem scénario 2)
- Hors référence IAS
- Sauvegarde sur VTL locale en rétention courte (plus ou moins 3 jours selon la taille du tampon disque) / Migration sur bandes distantes via des liens DWDM
- Performance de compression de 3 :1

Le scénario 4 de sauvegarder toutes les données sur DXi présente l'inconvénient :

- Nécessite l'acquisition de matériels DXi

Mais apporte les avantages suivants :

- Mutualisé IAS
- Sauvegarde sur VTL locale en rétention longue / réplication sur VTL distante
- Bénéfice de la déduplication : performance moyenne de 7 :1

J'ai également mené une étude financière afin de chiffrer ces différents scénarios. Le résultat ne peut être divulgué pour des raisons de confidentialité.

En prenant en compte le besoin de mutualisation avec les IAS, la solution VTL retenue fut la solution Quantum DXi, celle-ci se trouvant déjà en production dans sa version 5500 et donc maîtrisée par les exploitants IAS. De plus, elle a l'avantage de proposer une solution de **sauvegarde exclusivement sur disques** moins coûteuse (plus de nécessité de bandes, en particulier pour l'externalisation) grâce aux fonctionnalités de **déduplication** et de **réplication** des données entre systèmes distants (*pour de plus précision sur la solution DXi se reporter à l'annexe 2*).

Par contre, le DXi 5500 propose une volumétrie utile maximale de 10,8To et un débit de 800Go/h. Compte tenu du débit nécessaire (jusqu'à 867 Go/h) et de la volumétrie nécessaire (14,4 To par site), le choix s'est porté sur la solution supérieure dans la gamme au DXi 5500, c'est-à-dire le DXi 7500, lequel propose une volumétrie utile de 18 To utile et un débit d'1 To/h. Celui-ci a dû être **qualifié** afin de vérifier qu'il n'y avait pas de régression par rapport au DXi 5500.

Nous avons ainsi été les premiers clients de Quantum à tester la version bêta du produit DXi 7500. Nos retours sur le produit ont contribué à la version General Availability (GA).

b) SAN

J'ai évalué les besoins suivants pour le SAN :

- 2 ports SAN par media serveur (d'infrastructure) en double-attachement
- 1 port SAN par SAN media serveur (applicatif) en simple-attachement
- 4 ports SAN par DXi

J'ai finalement **exprimé les besoins suivants auprès de la MOE SAN** dans le cadre de ce projet :

- ports SAN pour les medias serveurs : 2*3 ports sur Aubervilliers ; 2*10 ports sur Vélizy

- ports SAN pour les SAN medias serveurs : 15 ports sur Aubervilliers ; 30 ports sur Vélizy
- ports SAN pour les DXi : 4 ports sur Aubervilliers ; 4 ports sur Vélizy

soit au total 25 ports sur Aubervilliers et 54 ports sur Vélizy.

En collaboration avec l'architecte SAN, nous avons choisi des **switches Brocade 4900 64 ports** affichant un débit de 4Gb/s avec une licence 32 ports sur Aubervilliers et 64 ports sur Vélizy.

Sur Aubervilliers, le switch Brocade 4900 sera installé et interconnecté au switch Brocade 4100 existant.

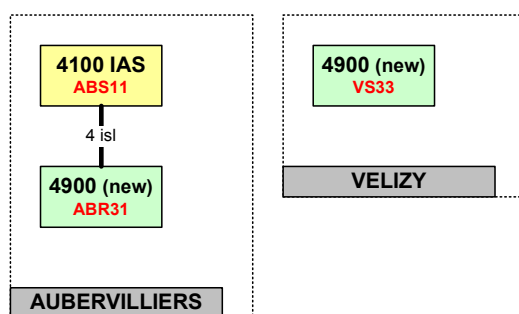


Illustration 27 : évolution des fabricas de l'IAS

c) serveurs

Le projet **Decommissionning** (également du programme MIO) m'a demandé de privilégier la réallocation de serveurs décommissionnés. Il m'a ainsi été proposé, entre autres, des serveurs Fujitsu PrimePower 450 et des IBM x445. Ces 2 modèles ont été largement éprouvés chez France Telecom en tant que serveurs de sauvegarde ; de plus, ma bonne connaissance de ce matériel constitua un atout lors de leur déploiement.

J'ai **validé la réutilisation de 4 PrimePower 450**, qui seront affectés comme medias serveurs de la zone trusted, car celle-ci héberge beaucoup de bases de données, sauvegardées via le SAN, et donc peu de clients IP ; une faible capacité disque pour le disk staging se révèle donc suffisante dans cette DMZ.

Par contre, aucune baie de disques locale n'est disponible, ni dans le périmètre du projet **Decommissionning**, ni au catalogue du constructeur. En effet, les baies EXP400 en connexion SCSI ont été retirées du catalogue. La qualification des x445 avec la seule carte SAS au format PCI-X, à savoir le modèle 4800 de chez Adaptec, ayant échoué, ce modèle n'a pas été retenu.

J'ai **validé la qualification du x3950**, successeur du x445 chez le constructeur IBM, ainsi que de sa baie de disques EXP3000 en connexion SAS, qui permettra d'utiliser la fonctionnalité de disk staging sur les medias serveurs desservant les clients IP des zones privée et publique. Au final,

j'ai exprimé le besoin d'un x3950 par zone privée/administration, et par zone publique, plus un pour l'IAS de Vélizy, plus le master soit 10 IBM x3950.

Ce modèle répond bien aux préconisations de l'éditeur [2] :

- 2 processeurs quad-core
- 4 Go de RAM (soit 2 GB par CPU) pour les masters et 8 Go pour les medias serveurs
- 2 ports Gigabit Ethernet (intégrés à la carte mère)

Calcul du débit réseau nécessaire pour les medias serveurs :

Afin de dimensionner l'architecture réseau, on admettra pour les calculs suivants que :

- 15% des données varient chaque jour
- une DMZ accueille au maximum 150 clients IP
- $sc = 150 * 5 \text{ Go} = 750 \text{ Go}$ pour chaque DMZ
- $dc = 150 * 3 \text{ Go} = 450 \text{ Go}$ pour chaque DMZ

Débit réseau nécessaire pour une DMZ :

- le samedi de 0h à 8h ou le dimanche de 0h à 8h :

$$\begin{aligned} & (sc/8 + 7/8*sc*0,15 + dc/2 + dc/2*0,15) / 8 \\ & = (0,25625*sc + 0,575*dc) / 8 \\ & = 56 \text{ Go/h} \\ & = 16 \text{ Mo/s} \end{aligned}$$

- tous les jours (sauf le jour de la totale) de 0 à 8h : $(dc*0,15) / 8 = 8,4 \text{ Go/h} = 2,4 \text{ Mo/s}$

Le réseau de la plate-forme MDSP étant en 100 Mb/s, une seule connexion ne serait pas suffisante, (surtout si l'on considère qu'en pratique une connexion réseau ne peut être utilisée à plus de 70% de sa bande passante théorique). J'ai donc validé avec les architectes réseaux l'utilisation de l'**agrégation de liens** afin d'obtenir une bande passante suffisante.

Un media serveur pourra ainsi absorber en entrée un débit théorique 200 Mb/s soit un débit utile de $200*70\% = 17,5 \text{ Mo/s}$. En considérant un débit moyen de sauvegarde d'un client IP égal à 3,5 Mo/s, un media serveur pourra sauvegarder $17,5/3,5 = 5$ clients IP en même temps. Ceci sera configuré en fixant à 5 le «maximum concurrent write drives» des Disk STU des medias serveurs afin d'assurer ce débit par client (afin de s'assurer qu'on ne peut pas se retrouver dans la situation d'essayer de sauvegarder 150 clients en même temps à 0,1 Mo/s chacun !) et de déclarer la même «backup window» (fenêtre de sauvegarde) du samedi et du dimanche pour tous les clients.

L'exploitant n'étant pas familiarisé avec l'agrégation de liens, il a fallu lui fournir les documentations afin de le mettre en place sur les serveurs (bonding mode 4) et sur les switches (trunking).

Nombre de Virtual Tape Drive (VTD) à affecter :

On affectera par défaut 2 VTD par SAN media serveur afin de sauvegarder ses propres données. Ce chiffre pourra être revu à la hausse après concertation avec les DBA en fonction du nombre de channels qu'ils souhaitent utiliser (on évite en effet de laisser un channel attendre la fin d'écriture d'un autre channel, avec le risque qu'il tombe en time-out). Il est en outre inutile d'ajouter un VTD pour la restauration. La plupart des logiciels de sauvegarde (dont NetBackup) affectent une priorité plus grande aux restaurations par rapport aux sauvegardes ; si une restauration devait intervenir pendant une sauvegarde en cours, on désactiverait au besoin cette sauvegarde le temps de la réouverture du service. Le «maximum concurrent write drives» de la STU sera fixé au nombre de VTD affectés.

On affectera par défaut 3 VTD par media serveur afin de destager les données des clients IP sur le DXi, le troisième se justifiant pour la tolérance de pannes (si un VTD tombe en erreur et mis «down» par NetBackup) et pour la restauration (sinon il faudra attendre la fin du(des) job(s) de sauvegarde en cours afin de pouvoir lancer une restauration, laquelle est généralement prioritaire). Ceci sera configuré facilement sous NBU en limitant à 2 «maximum concurrent write drives» les STU des medias serveurs, soit le nombre de VTD affectés moins un.


d) espace catalogue des master serveurs

Le master d'Aubervilliers, un IBM x3850, étant déjà en production, je me suis assuré qu'il pourrait **absorber la charge supplémentaire** de 262 clients IP, 15 SAN medias serveurs et 3 medias serveurs.

Afin de conserver une cohérence et une exploitation simplifiée, le master de Vélizy sera également un IBM x3950, mais avec moins de mémoire qu'un media serveur (cf. les préconisations éditeurs ci-dessus).



Site	Catalogue	Serveur	Nb d'objets
Aubervilliers	msptina04	msepap738	2 320 571
	msptina05	msepap738	6 426 875
	msptina06	msepap738	10 299
	msptina10	msepap738	1 133 573
	msptina11	msepap738	3 308 329
Velizy	msitina01	msepva730	10 621 856
	msqtina01	msepva730	2 080 766
	msqtina02	msepva730	6 417 037
	msqtina03	msepva730	4 991 572
	msqtina04	msepva748	4 436 182
	msqtina05	msepva748	5 259 034
	msqtina06	msepva748	6 794 626
	msptina01	msepva732	2 190 567
	msptina02	msepva732	5 626 843
	msptina03	msepva732	14 392 229
	msptina07	msepva732	1 209 671
	msptina08	msepva732	2 928 254
	msptina09	msepva760	3 819 753
	msptina12	msepva774	817 054
			84 785 091

Tableau XI : taille des 19 catalogues TiNa

Calcul de la volumétrie nécessaire pour héberger le catalogue sur les masters  :

Sans TIR : NetBackup catalog size = 120 * (number of files)

Avec TIR : NetBackup catalog size = 2 * 120 * (number of files)

Le True Image Recovery with Move Detection (TIR) permet de tenir compte des fichiers supprimés, déplacés, ou extraits d'une archive. Cette fonctionnalité, absente sous TiNa , permet de retrouver l'état exact de la machine sauvegardée avec des sauvegardes incrémentales. Cette fonctionnalité, proposée par la MOE  et approuvée par les utilisateurs, sera donc implémentée. Il en sera de même pour d'autres fonctionnalités, comme le Checkpoint Restart, qui permet de positionner des points de reprise à intervalles réguliers ou à la demande, permettant respectivement une reprise sur incident ou à la demande à partir du dernier point de reprise, au lieu de reprendre la sauvegarde depuis le début.

			Taille catalogue en Mo	
Serveur	Nb d'objets	Site	sans TIR	avec TIR
osiasa01	22 714 165	Aubervilliers	2 599	5 199
osiasv01	62 070 926	Velizy	7 103	14 207

Tableau XII : estimation de la taille des 2 catalogues NBU

Nous avons donc besoin de :

- 14,2 Go sur Vélizy à héberger sur les disques internes de 73 Go (en RAID 1), sachant que le système et les applications occuperont environ 20 Go, donc le catalogue pourra être hébergé sur un filesystem d'une cinquantaine de Go qui sera occupé à 30% environ en fin de projet.
- 5,25 Go supplémentaires sur Aubervilliers à héberger sur les disques internes de 73 Go (en RAID 1), sachant que le système et les applications occupent environ 20 Go, et que le catalogue occupe aujourd'hui 5,1 Go sur un filesystem de 49 Go. On utilisera donc à la cible $5,1 + 5,25 = 10,35$ Go soit environ 20% de l'espace disponible.

Nous n'avons donc pas besoin d'investir dans une coûteuse baie de disques pour les masters. En effet, la capacité des disques internes suffira pour héberger de manière sécurisée (Raid 1) le système, les produits NetBackup et Patrol, et le catalogue (même non compressé), tout en permettant d'absorber de nouveaux clients. De plus, les I/O, limités à l'écriture d'une entrée (120 octets) par fichier sauvegardé, sont facilement absorbables, même par des disques internes de serveur.

Nous restons bien en-deçà des limites préconisées par Symantec[3] : taille du catalogue en ligne inférieure à 750 Go et nombre d'entrées par catalogue inférieur à 1 000 000.

6.4. Intégration dans l'IAS

Pour la définition de l'architecture cible, il a fallu prendre en compte les éléments suivants :

- **Conservation de l'architecture TiNa** à des fins de restauration.
- **Respecter les règles de sécurité** réseau des nids de l'infrastructure MDSP.
- **Maintenir les performances** de sauvegardes/restaurations.
- **Ne pas modifier les règles d'exploitation** des sauvegardes en IAS.
- **Libération de l'infrastructure TiNa** (serveurs, switches SAN et robotiques physiques) en fin de projet.

Pour des raisons de sécurité, il a donc été décidé d'utiliser les masters en IAS et de déployer dans chaque nid un media serveur afin de **limiter au maximum les interactions réseau** (cf. chapitre précédent). Les masters étant en IAS ceci a également permis de **réduire les coûts RH** en utilisant les équipes d'exploitation des sauvegardes IAS qui ont repris l'administration de l'infrastructure déployée.

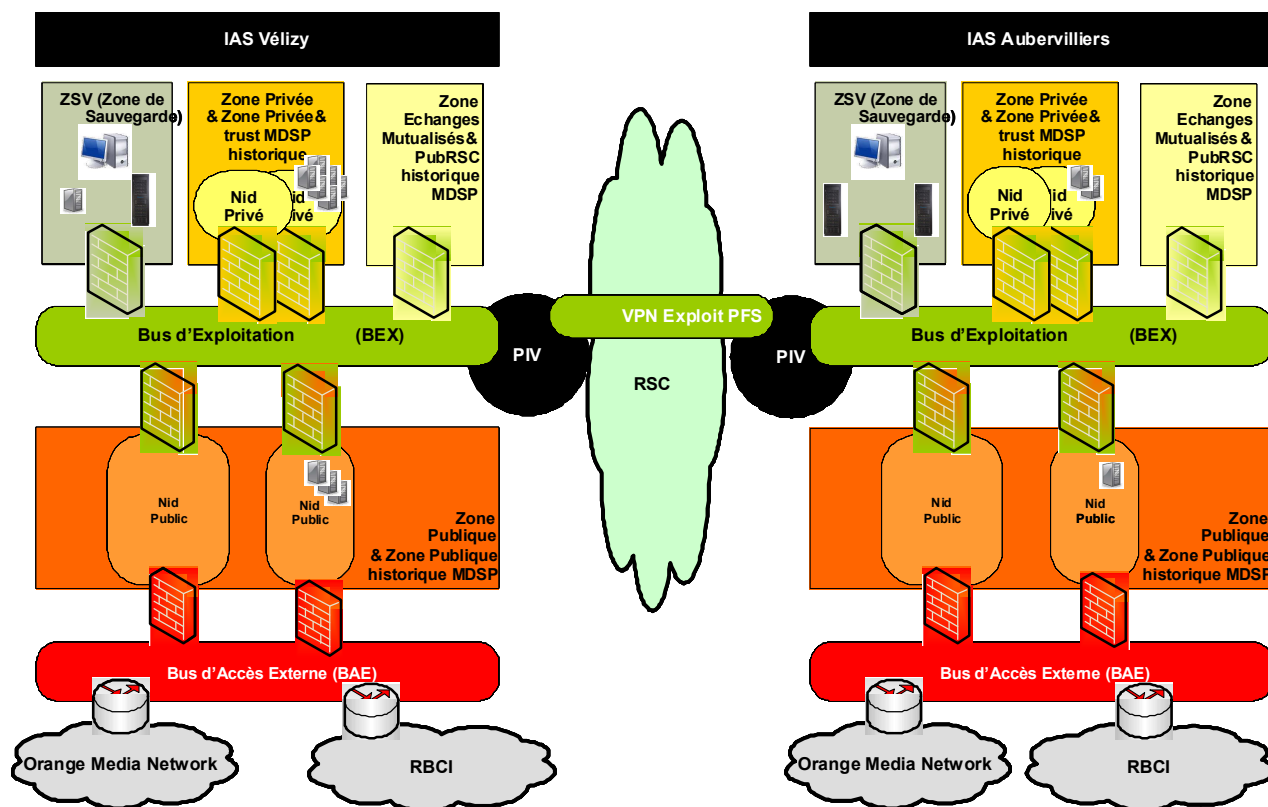


Illustration 28 : intégration dans le réseau IAS

J'avais initialement prévu de migrer les serveurs sous NBU au fur et à mesure de leur migration réseau (afin qu'ils puissent nativement communiquer avec l'infrastructure de sauvegarde). Les pré-requis au niveau des flux étaient alors les suivants :

- sauvegarde Tina : internes au nid, à conserver pour le retour arrière et les restaurations sur Tina,
- sauvegarde NetBackup : internes au nid, et entre les nids et la ZSV,
- exploitation Tina : STP doit pouvoir continuer à exploiter Tina dans l'IAS,
- réplication DXi : sur la Passerelle Inter VPN (PIV) entre les 2 IAS.

En cours de projet j'ai dû **revoir les interconnexions SAN et réseau**, la solution initialement prévue de migrer les architectures MDSP en IAS n'étant **économiquement pas viable**. Le projet de migration dans l'IAS a en effet conclu le 26 septembre 2008 de :

- Migrer au cas par cas, des services pérennes de MDSP dans les IAS lors d'une nouvelle version majeure ou d'une refonte de l'architecture et ceci en fonction du ROI individuel.
- Tout nouveau service Mobile sera construit en IAS.
- Le périmètre du projet est réduit à l'interconnexion MDSP-IAS.
- Les chantiers additionnels à mener sont de rationner et mutualiser les outils d'exploitation (notamment outils de supervision) entre MDSP et IAS, et d'étudier la réduction du nombre d'environnements (notamment les environnements de pré-production) et du nombre d'équipements réseau (décommissionnement).

a) réseau IP

Pour ne pas dépendre du projet MIO de migration Réseau de MDSP vers les IAS (consolidation des plates-formes MDSP et IAS), le scénario de raccordement réseau des infrastructures hardware déployées en IAS Vélizy (salle VS33) et en IAS Auber (salle ABS31) doit :

- permettre de poursuivre le déploiement de la nouvelle infrastructure de Sauvegarde indépendamment du projet de migration réseau MDSP vers IAS
- permettre de démarrer l'activation de l'agent NBU sur les clients MDSP indépendamment du projet de migration réseau MDSP vers IAS
- permettre de répondre au besoin de sauvegarde avec NBU (au lieu de TiNa) pour les nouveaux projets IAS et MDSP.

Je rappelle que le master doit pouvoir communiquer avec tous les clients de son site et chaque media serveur doit pouvoir communiquer avec les clients dont il met les données sur bande, c'est-à-dire aux clients de sa DMZ dans l'architecture que j'ai proposée.

Les deux réseaux Ethernet MDSP et IAS étant cloisonnés et d'architecture différente (VLAN d'administration et de service dédiés par outil en IAS, VLAN d'administration et de production par DMZ sur MDSP), la solution la plus simple pour que le master puisse communiquer avec ses clients consiste à utiliser le LAN d'administration de la plate-forme. Ce réseau d'administration comporte un VLAN global à la plate-forme qui permet d'accéder à la patte d'administration de l'ensemble des serveurs de la plate-forme. On retrouve ainsi dans ce VLAN les bastions, les serveurs de supervision... et donc à terme le master serveur NBU.

L'architecture cible, qu'il a fallu faire valider auprès des architectes sécurité, propose donc **un triple attachement réseau Ethernet des masters** :

- un lien côté MDSP : sur le réseau d'administration MDSP, dans le VLAN d'administration global, afin de pouvoir sauvegarder les clients de la plate-forme MDSP. Ce lien permet également d'administrer le serveur depuis le bastion MDSP, en cas de problème sur TDIMG.
- un lien côté administration IAS dans la ZSV, dans le VLAN administration sauvegarde IAS afin de permettre l'administration du master par les exploitants sauvegarde après identification sur le bastion TDIMG.
- un lien côté service IAS dans la ZSV, dans le VLAN service sauvegarde IAS afin de sauvegarder les clients de l'IAS.

Remarque : une zone IAS demeure une notion logique.

b) réseau SAN

Le nouveau DXi d'Aubervilliers doit pouvoir sauvegarder les medias serveurs :

- de production MDSP (d'Aubervilliers),
- de l'infrastructure NBU en cours de déploiement.

Le DXi de Vélizy doit pouvoir sauvegarder les medias serveurs :

- de production MDSP (de Vélizy),
- de pré-production,
- de l'infrastructure NBU en cours de déploiement,
- de l'IAS de Vélizy.

Je rappelle que le SAN Backup Orange doit être abandonné pour des raisons de fin de maintenance et d'exploitation (*cf. chapitre 6.3.c*).

J'ai discuté les scénarii suivants avec l'équipe SAN :

Scénario b1 : Interconnexion des SAN de backups Orange et IAS

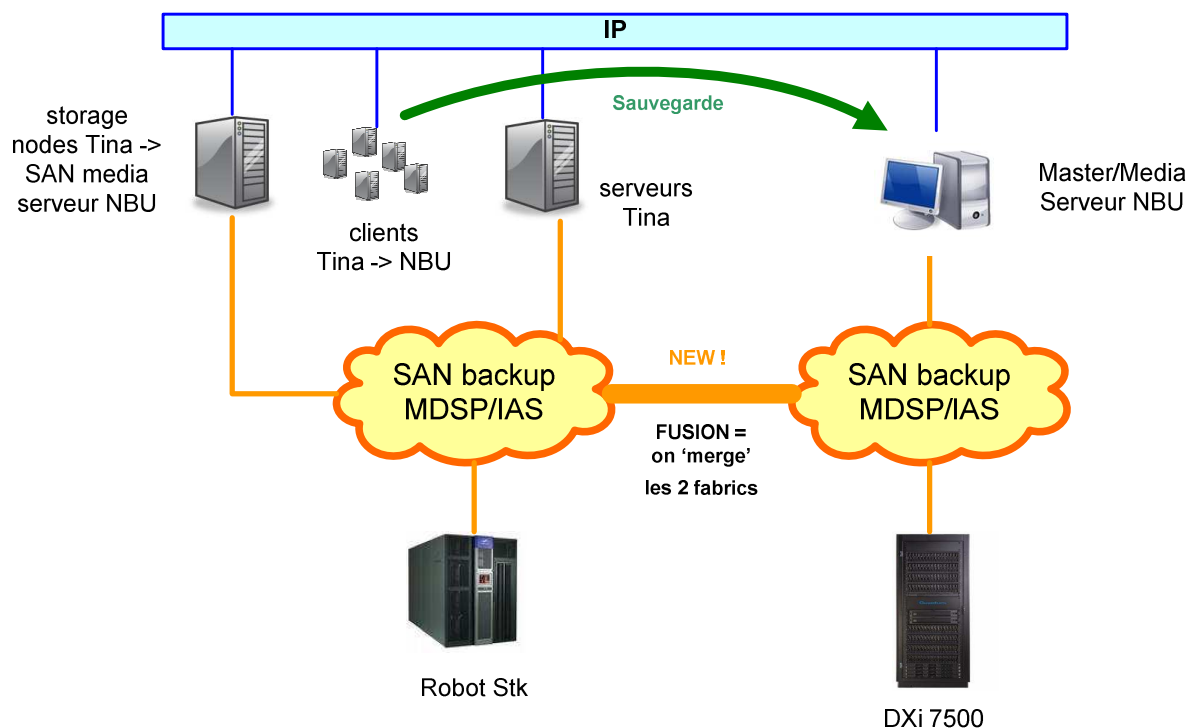


Illustration 29 : scénario 1 d'architecture SAN

- Dans ce cas de figure, on interconnecte les SAN de Backup Orange et IAS. Les storage nodes Tina accèdent au DXi connecté au SAN de backup MDSP/IAS via l'interconnexion mise en place. Les restaurations depuis le robot STK restent possibles sans aucune modification côté SAN backup Orange.
- **Actions à prévoir (au niveau du SAN) :**
 - Si non disponibles, tirer des rocade optiques temporaires entre un switch Orange et un switch IAS.
 - Interconnecter les deux SAN de backup, fusionner («merger») les fabrics, prendre en compte la différence de stratégie de zoning entre l'environnement MDSP et IAS.
 - Zoner les storage nodes avec le DXi.
- **Avantages :**
 - Simple à mettre en œuvre, mais on se contente de réaliser physiquement l'interconnexion sans homogénéiser les stratégies de zoning.
 - Nécessite peu de fibres optiques (2 ou 4 paires de fibres maximum).
- **Inconvénients :**
 - Le SAN de Backup Orange est un SAN mutualisé qui englobe le périmètre MDSP, mais héberge également d'autres applications hors périmètre, ce qui pose le problème de la responsabilité au niveau de l'exploitation.
 - On fait du zoning par port côté Orange et du zoning par WWN côté IAS. L'interconnexion ne pose pas de problème technique particulier et les deux types de zones peuvent cohabiter au sein de la même fabric. En revanche, il y a risque de créer des zones mixtes (WWN + port) lesquelles sont fortement déconseillées par le constructeur Brocade.
 - Gérer à la fois du zoning par port et du zoning par WWN au sein de la même fabric complexifie l'exploitation et l'administration.

Scénario b2 : Double attachement des Storage Nodes Tina

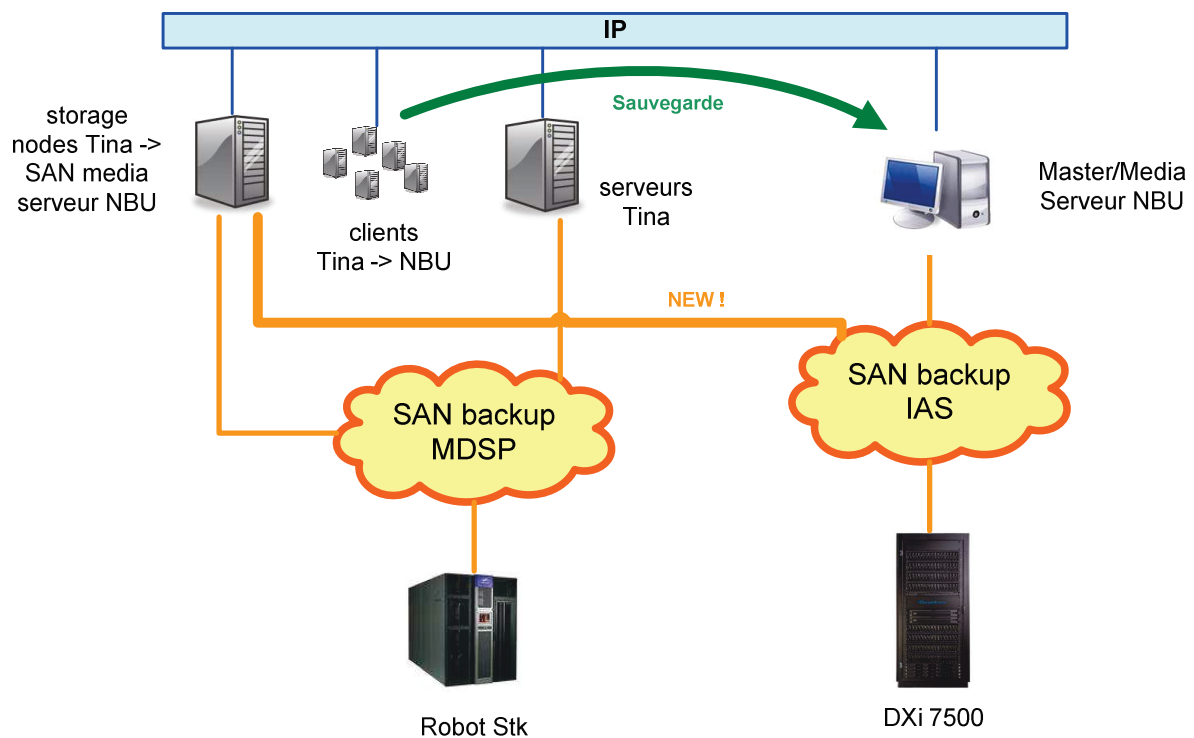


Illustration 30 : scénario 2 d'architecture SAN

- Dans ce cas de figure, on interconnecte les storage nodes à la fois au SAN de Backup Orange et au SAN de backup IAS. Les storage nodes Tina convertis en SAN media serveurs NBU accèdent à la fois au robot pour assurer les restaurations et au DXi pour assurer les 'nouvelles' sauvegardes.
- **Actions à prévoir (au niveau du SAN) :**
 - S'assurer que tous les storage nodes possèdent au moins 1 port HBA non utilisé.
 - Localiser physiquement chaque storage node sur chacun des sites (salle, dalle, baie).
 - Tirer autant de fibre que nécessaire entre tous les storage nodes et le switch 4900 sur le SAN de backup IAS.
 - Zoner les storage nodes avec le DXi.
- **Avantages :**
 - Il n'y a pas d'interconnexion entre les SAN de backup Orange et IAS (domaine de responsabilité bien délimité).
 - Aucune modification de configuration (zoning) à prévoir côté MDSP.
 - Modification de configuration (zoning) basique à prévoir côté IAS.
- **Inconvénients :**
 - Chaque storage node doit posséder au moins 1 port HBA non utilisé pour réaliser le double attachement (au minimum : carte HBA bi-port ou deux cartes HBA mono-port). Certains serveurs ne remplissent pas ce pré-requis.
 - Pour chaque storage node, il faudra prévoir une rocade optique entre la salle où est implanté le serveur et la salle où sera présent le SAN IAS, ce qui nécessite donc un câblage important. Dans tous les cas ce câblage sera nécessaire en fin de phase transitoire

(plus de connexion au SAN de backup MDSP) mais le temps nécessaire à la mise en place de ce double attachement risque de faire prendre du retard au projet.

- Consommation de ports SAN importante (mais temporaire ; ceci ne constitue pas réellement un problème actuellement car il y a suffisamment de ports disponibles).

Scénario b3 : Double attachement des DXi

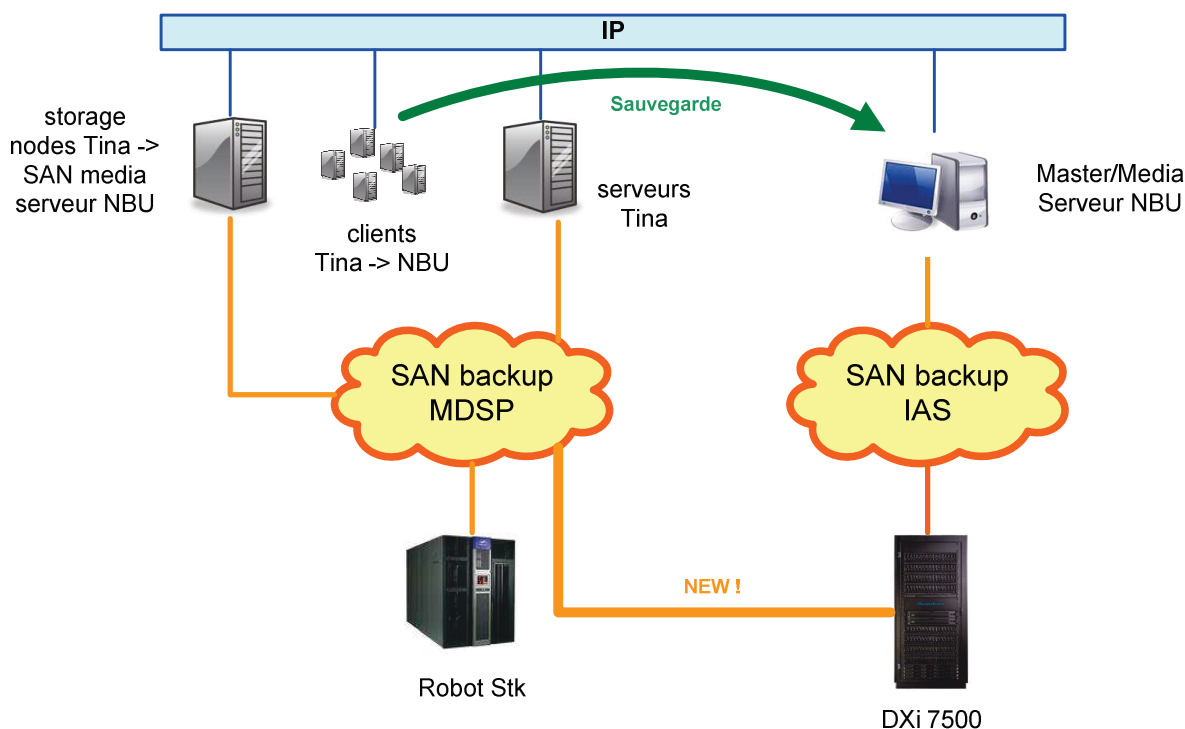


Illustration 31 : scénario 3 d'architecture SAN

- Dans ce cas de figure, on interconnecte le DXi à la fois au SAN de backup Orange et au SAN de backup IAS. Les storage nodes Tina convertis en SAN media serveurs NBU accèdent à la fois au robot pour assurer les restaurations TiNa et au DXi pour assurer les sauvegardes NBU.
- **Actions à prévoir (au niveau du SAN) :**
 - Câbler le DXi sur un switch SAN Orange via 2 paires de fibres optiques. Au besoin, faire tirer les fibres entre les salles.
 - Zoner les storage nodes avec le DXi sur le SAN de backup Orange.
- **Avantages :**
 - Il n'y a pas d'interconnexion entre les SAN de backup Orange et IAS (domaine de responsabilité bien délimité).
 - Aucune modification de configuration (zoning) à prévoir côté IAS.
 - Modification de configuration (zoning) basique à prévoir côté MDSP.
 - Très peu de câblage (ou de tirage de fibres optiques) à réaliser.
- **Inconvénients :**
 - Mettre en place du persistant binding au niveau des storage nodes convertis en SAN media serveurs.

Nous avons choisi le scénario 3 qui consiste en un **attachement SAN multiple des systèmes DXi** sur les SAN de backup Orange et IAS (double attachement sur Aubervilliers et triple attachement sur Vélizy pour prendre en compte le SAN de pré-production MDSP). Il s'agit de l'architecture qui présente le moins de contraintes et qui sera la plus rapide à déployer.

On connectera donc le DXi d'Aubervilliers :

- temporairement au SAN Backup Orange et au SAN Backup IAS,
- à terme uniquement au SAN Backup IAS.

et le DXi de Vélizy :

- temporairement au SAN Backup Orange, au SAN de pré-production et au SAN Backup IAS,
- à terme au SAN de pré-production et au SAN Backup IAS.

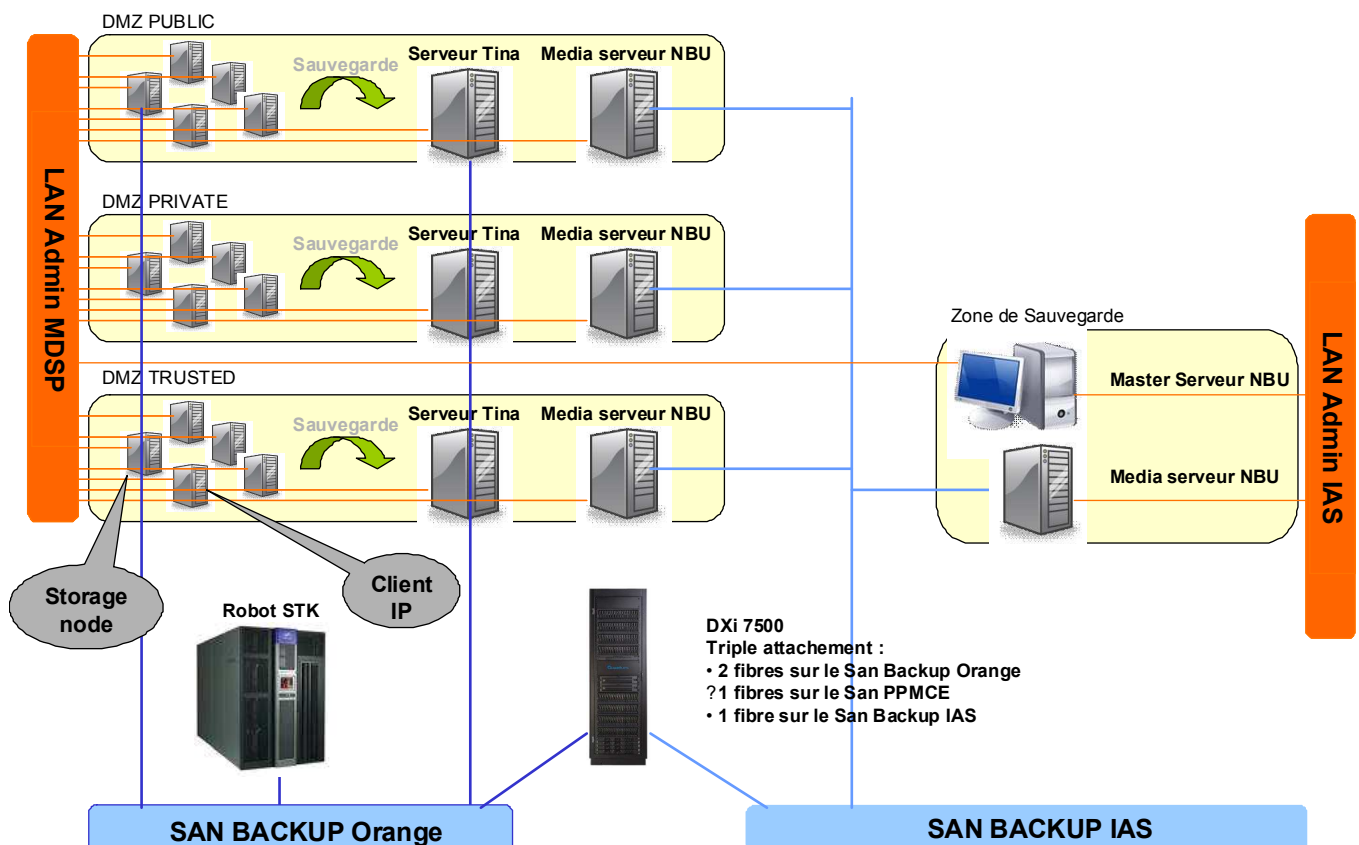


Illustration 32 : sur Vélizy, double attachement réseau d'administration du master et triple attachement du DXi

6.5. Livrables

Les livrables réalisés lors de ce stage furent les suivants :

- Dossier d'architecture Technique (DAT) (*voir chapitre 6.6*),
- Scénario de migration (*voir chapitre 6.9*),
- Scénario de sauvegarde des bases de données (*voir chapitre 6.7*),
- Étude de coût de l'infrastructure, incluant une étude comparative de l'architecture sauvegarde sur disques versus architecture bandes *partiellement discuté dans ce mémoire (chapitre 6.3.a) pour des raisons de confidentialité sur les tarifs accordés à la société France Telecom par ses fournisseurs*,
- Plan Management Projet (PMP) (*voir chapitre 5.6*),
- Documents de passage de jalons,
- Procédures de migration et de suppression de l'infrastructure TINA,
- Kit de communication sur le nouveau mode de sauvegarde des bases de données.

Lors de ma mission, j'ai également dû assurer le **soutien** au quotidien auprès de l'exploitation, et un **suivi hebdomadaire** auprès de l'ensemble des acteurs et de l'équipe projet transverse.

6.6. Rédaction du DAT et passage en PCVAT/CVAT

J'ai dû définir tous les éléments de l'architecture dans un Dossier d'Architecture Technique (DAT) que j'ai défendu en Comité de Validation des Architectures techniques (CVAT📖) (*cf. présentation du CVAT au chapitre 4.2.a*).

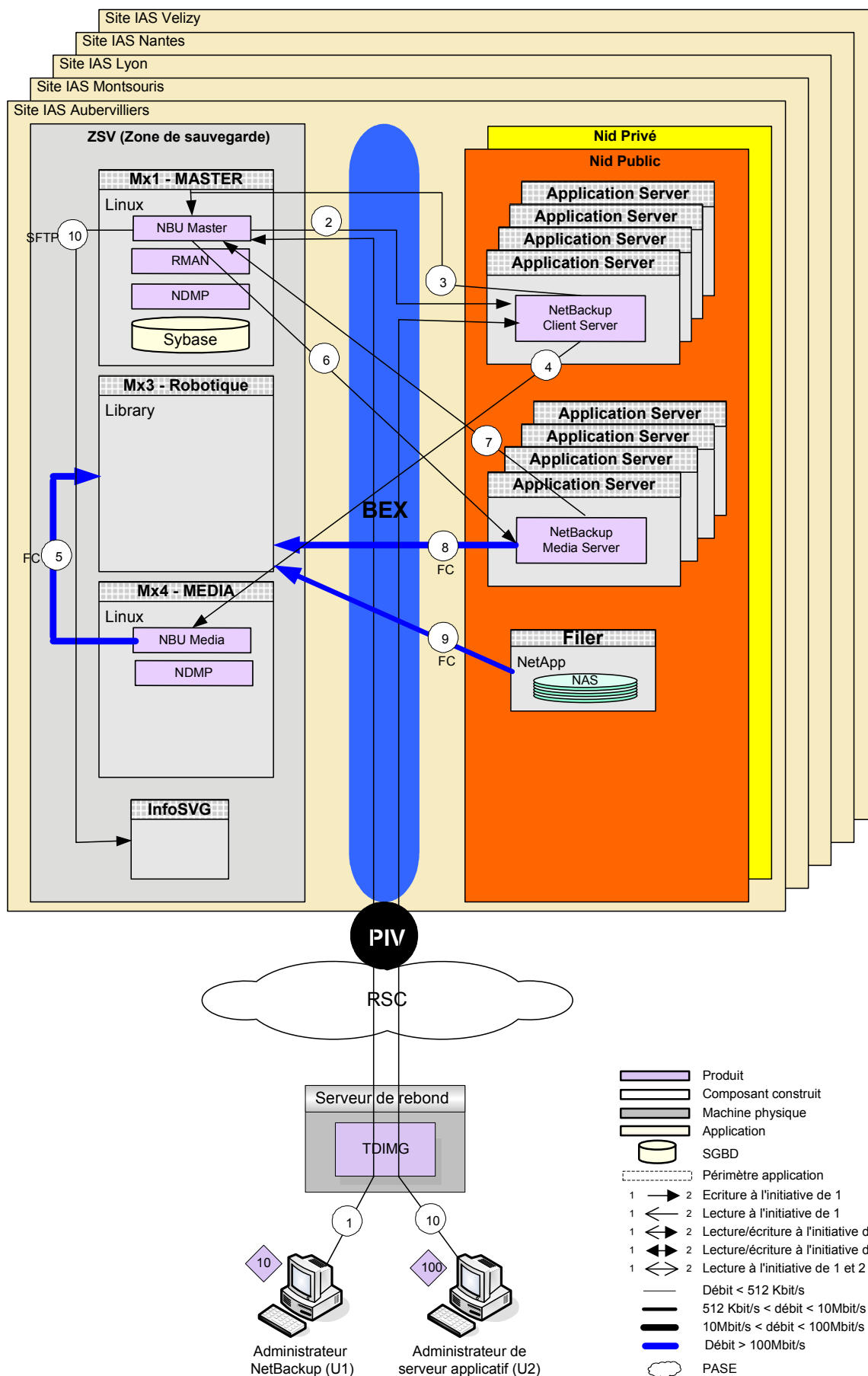


Illustration 33 : schéma d'architecture

On m'a demandé de compléter le DAT actuel (version G1R0) présentant l'architecture NetBackup dans les IAS, afin de n'avoir à maintenir qu'un seul DAT, en limitant les spécificités liées à mon projet.

J'ai ainsi décrit les **évolutions** suivantes dans la version G1R1 :

- l'augmentation des clients et de la volumétrie,
- la prise en compte du nouvel IAS de Vélizy,
- les serveurs NetBackup ajoutés (le master de Vélizy et les 13 medias serveurs),
- la configuration des matériels x3950, PrimePower 450 et DXi 7500 et leur scalabilité,
- l'agent de sauvegarde des bases de données DB2.

J'en ai également profité pour mettre à jour la G1R0 avec les éléments suivants :

- présentation des options TIR, Checkpoint Restart et Multiple Data Streams,
- ajout des prescriptions (Sauvegardes système, pas de backup selection "/", exclusions par défaut, rétentions standards),
- changement des interlocuteurs (MOA, équipe d'exploitation),
- diverses corrections mineures.

Items évolutions \ Versions	G1R0	G1R1
Nouveaux flux		non
Augmentation volumétrie		oui (ajout des clients MDSP) soit 26 To
Rajout d'IHM		non
Rajout de fonctionnalités		non
Upgrade CPU		non
Nombre utilisateurs		
Mise en œuvre PRA		non
Changement version produits/OS		non
Achat/Rajout de serveurs		achat 10 serveurs + 2 DXi 7500 ajout 4 serveurs décommissionnés
Ouverture sur internet		non
Date passage en PCVAT		10/04/08
Date passage en CVAT		15/05/08
Jalon de validation de l'analyse d'impacts		T1

Tableau XIII : justification du passage en CVAT

Lors du passage en CVAT📖, les architectes ont **constaté** l'architecture proposée. Ils ont en particulier relevé les points forts suivants :

- **Scalabilité**📖 **horizontale** permettant de répondre avec la même qualité de service en cas de montée en charge.
- L'agent **NDMP** (option) qui permet de sauvegarder directement les données des filers sans faire passer le flux des données sur le réseau.
- **Virtualisation robotique** permettant de s'affranchir de problèmes mécaniques.

Ils ont néanmoins relevé les points faibles suivants de l'architecture :

- **Pas de backup pour le master**📖 serveur, bien que ce point soit conforme au cahier des charges.
- Utilisation de l'OS **Solaris 9**.

J'ai dû justifier ces 2 points faibles relevés comme suit :

- Le master📖 serveur constitue effectivement un SPOF (Single Point of Failure). Le **RTO**📖 **de l'infrastructure NBU est de 24h** ce qui correspond au délai pour réinstaller un master📖 NBU et restaurer le catalogue. Ceci suppose néanmoins qu'il ne s'agisse pas d'un problème hardware majeur, auquel cas il faudrait disposer d'un serveur de spare sur chaque site IAS📖, ce qui n'est pas envisageable économiquement. *Une des solutions qui pourra être étudiée serait d'installer le master dans une VM (solution supportée par l'éditeur Symantec à partir de la version 6.5.3).*
- Le produit NBU n'est **pas qualifié sous Solaris 10** (qualification prévue à partir de la version 6.5). L'OS Solaris 9 demeure largement éprouvé avec la solution master📖 NBU chez France Telecom.

J'ai également noté les recommandations du CVAT📖 auxquelles j'ai déjà pu apporter des réponses :

- R1 : Étudier, avec le projet IAS📖, la mise en œuvre d'une solution de **PRA**📖 permettant de répondre, si besoin, à un RTO📖 sur sinistre site. *Un **projet (distinct) de PRA**📖 de l'infrastructure NBU des IAS📖 sera lancé prochainement.*
- R2 : Étudier la migration en Linux **Red Hat 5** (Platon G8R0). *La migration vers Linux Red Hat 5 n'est pas à l'ordre du jour car le produit NetBackup en version 6.0 n'est **pas compatible** (compatibilité et qualification prévues à partir de la version 6.5).*
- R3 : Étudier la migration des serveurs Linux en hébergement **Ecocenter**📖. *La virtualisation sur Ecoserveur X86 n'est pas envisagée dans l'état actuel des choses car les medias serveurs ont besoin d'un **attachement SAN**📖/**FC**📖 (incompatible avec une machine virtuelle VMware).*

6.7. Sauvegarde des bases de données

a) Oracle

Le processus des sauvegardes de bases de données Oracle s'avère **lourd et inadapté** ; ces sauvegardes sont :

- **sécurisées** : **RPO proche de 0** grâce aux technologies de réplication de baies et d'espace de stockage BCV et SRDF,
- **sans impact** sur le serveur hébergeant la base de données (puisque la sauvegarde se retrouve déportée sur un autre serveur appelé «Infocentre»),
- **coûteuses** : besoin de 4 fois l'espace disque nécessaire et d'un serveur supplémentaire «Infocentre» sur le site distant (sur lequel s'exécute la sauvegarde RMAN),
- **complexes** à exploiter : son support nécessite des compétences en ordonnancement, Oracle, SAN et sauvegarde (compétences réparties dans 4 équipes d'exploitation distinctes),
- **différentes** entre la production et la pré-production, ce qui ne permet pas de reproduire certains comportements,
- au final **non justifiées**.

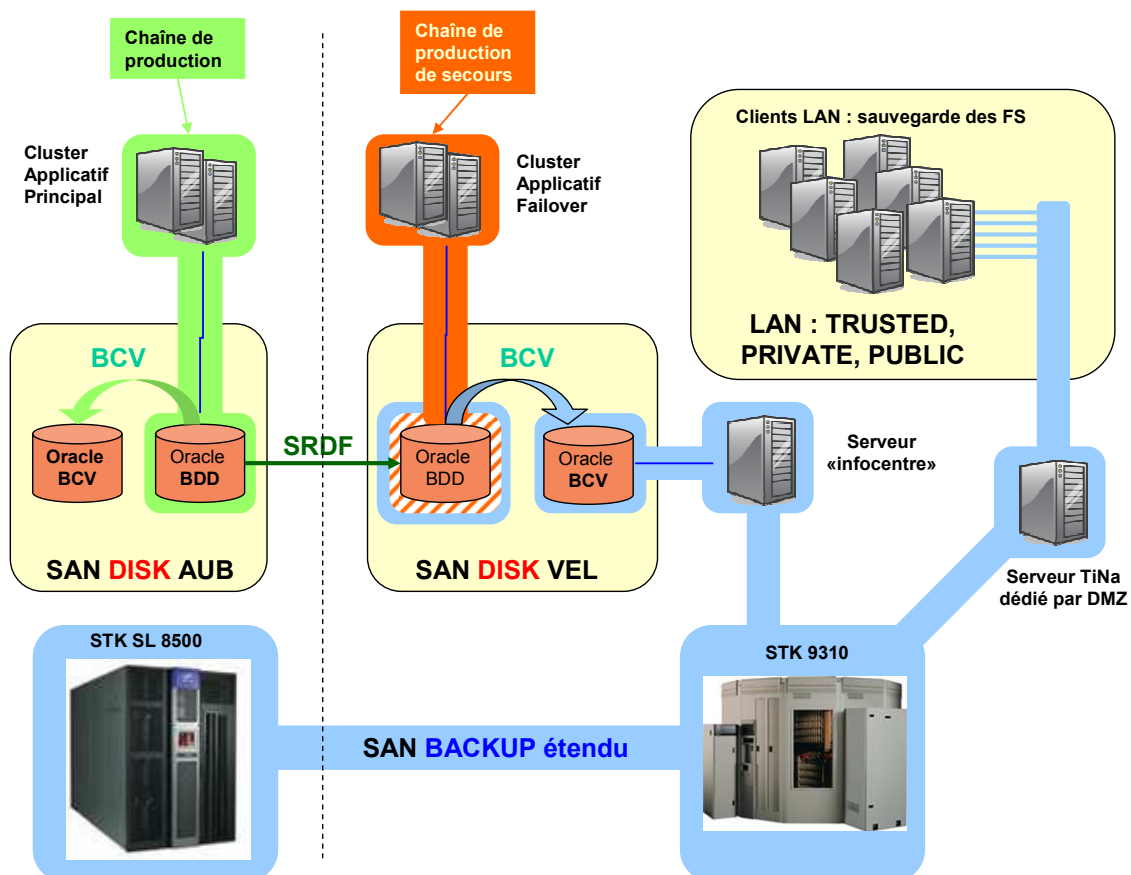


Illustration 34 : solution actuelle : RMAN Server Less + TiNa + STK

Le synopsis de l'ordonnancement actuel en production est le suivant :

- passage de la base en mode begin backup,
- BCV sur les 2 sites,
- split BCV sur Vélizy,
- montage du split BCV sur un nœud de stockage Tina dédié (également appelé «Infocentre»),
- sauvegarde via Tina + RMANOO sur robot STK depuis ce serveur «Infocentre».

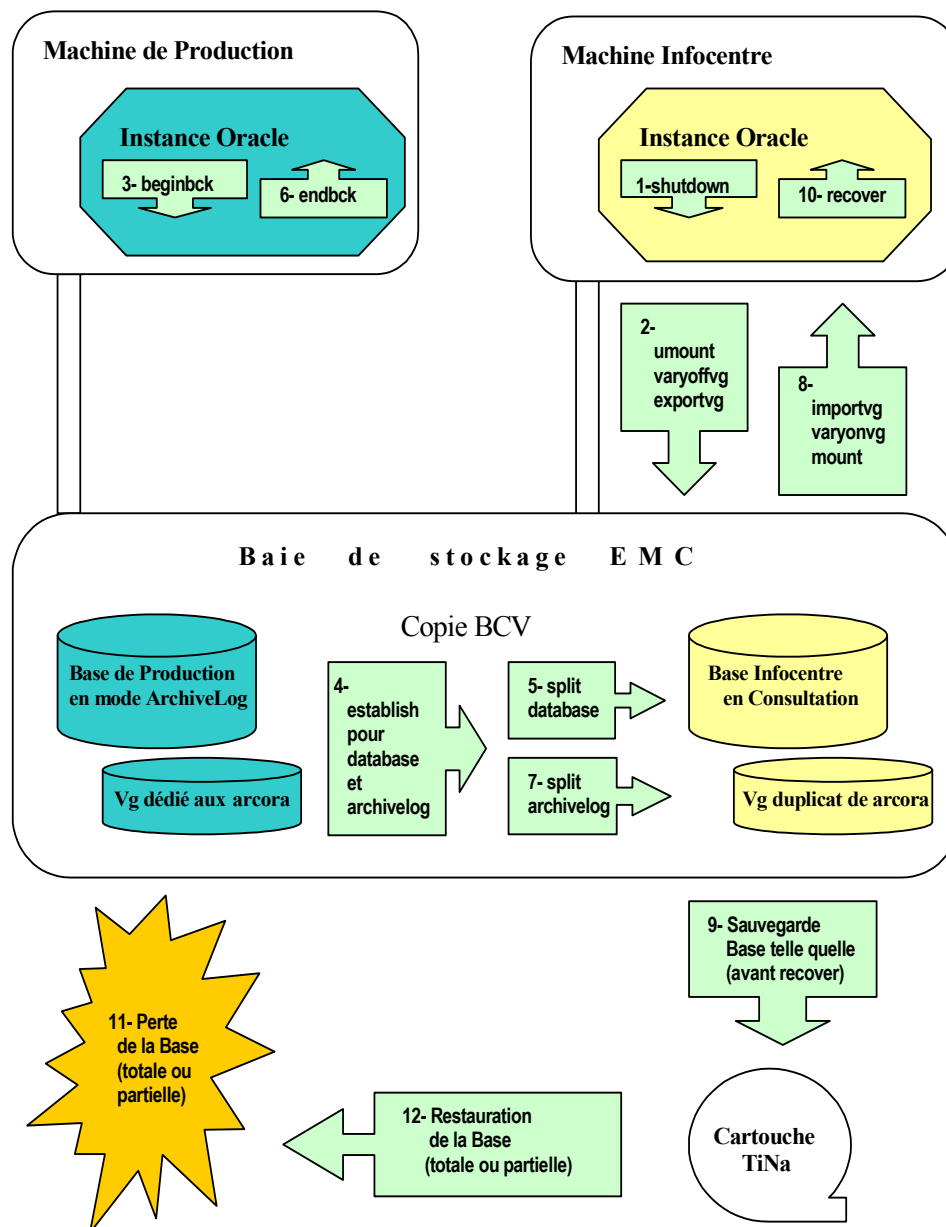


Illustration 35 : ordonnancement des sauvegardes Server Less

Je rappelle qu'un de mes objectifs est d'étudier la possibilité de **simplifier ce processus** de sauvegarde dans le cadre du projet de migration, en proposant une méthode **fiable** et **généralisée**.

J'ai ainsi **proposé et validé** avec les différents acteurs du projet de se défaire de ce processus à base de BCV📖, afin de **sauvegarder la base directement sur le serveur de production**. Un pilote permettra de vérifier le bon fonctionnement de ce nouveau processus et son faible impact sur les performances.

Ci-dessous une représentation et un comparatif des 2 méthodes :

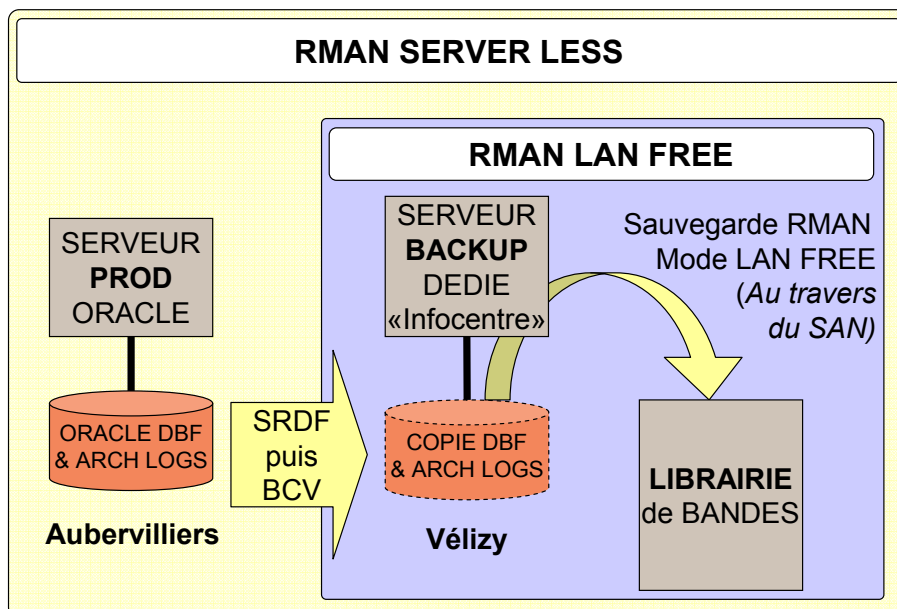


Illustration 36 : RMAN Server Less et RMAN Lan Free

SOLUTION GENERALISEE RMAN LAN FREE	SOLUTION ECARTEE RMAN SERVER LESS
Description : <ul style="list-style-type: none"> ➔ Mécanisme de sauvegarde à chaud (sans arrêt) ➔ Sauvegarde prise en charge par le serveur de production lui-même 	Description : <ul style="list-style-type: none"> ➔ Mécanisme de sauvegarde tiède (sans arrêt et sans impact sur la PROD) ➔ Sauvegarde des données prise en charge par un serveur tiers dédié à cette fonction ➔ <u>Basé sur les BCV</u>
Avantages : <ul style="list-style-type: none"> ➔ Solution de référence FT ➔ Implémentation rapide 	Avantages : <ul style="list-style-type: none"> ➔ Sauvegarde en parallèle de la PROD ➔ Solution historique sur MDSP
Inconvénients : <ul style="list-style-type: none"> ➔ Impact sur le serveur de PROD 	Inconvénients : <ul style="list-style-type: none"> ➔ Demultiplication des BCVs ➔ Implémentation lourde et complexe ➔ Création d'un pôle de compétences pour portage NBU / Oracle 10G

Tableau XIV : comparaison du mode de sauvegarde des bases Oracle

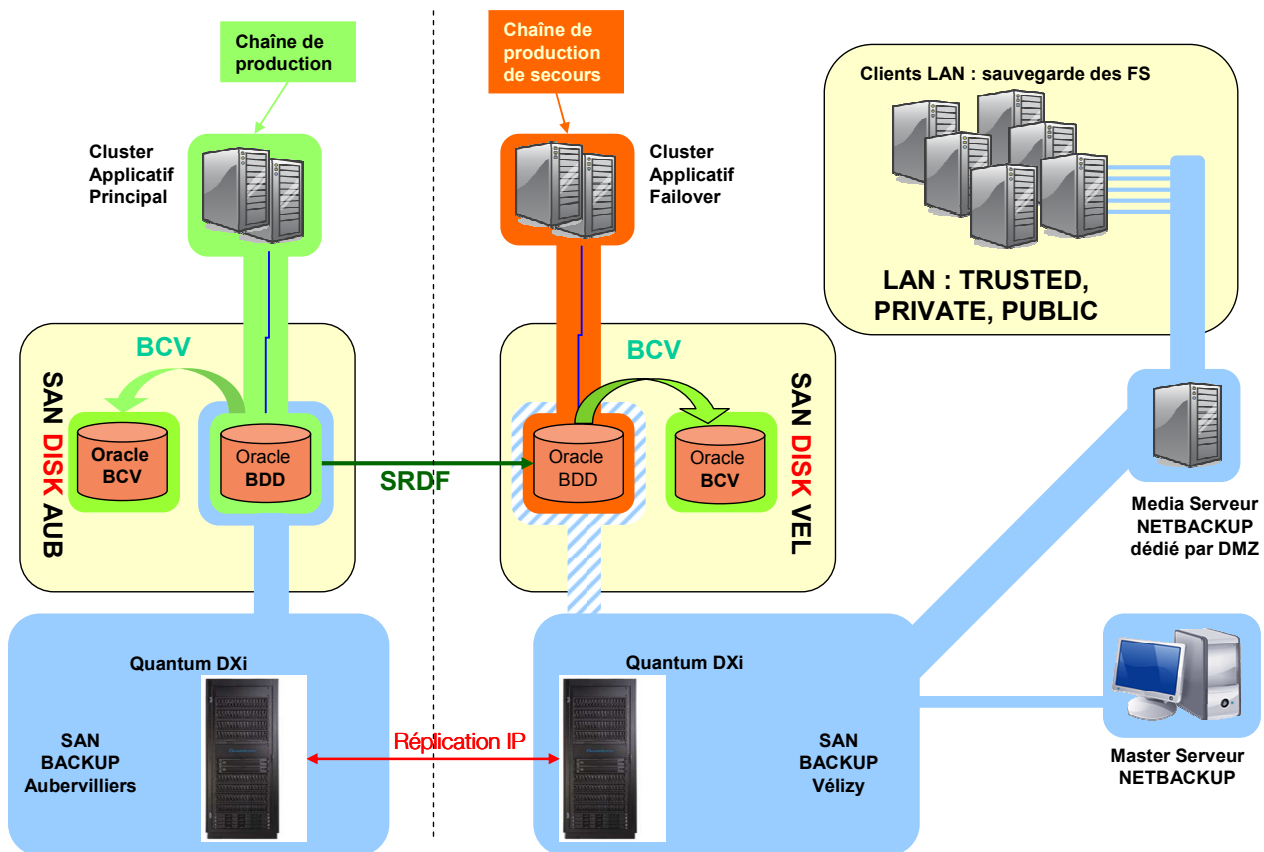


Illustration 37 : solution cible : RMAN LAN Free + NBU + DXi

J'ai fourni aux équipes en charge de l'outil d'ordonnancement \$U les modifications à apporter dans les UPROC de sauvegarde afin de respecter le synopsis d'ordonnancement suivant :

- passage de la base en mode begin backup,
- BCD sur les 2 sites,
- sauvegarde de la base via NBU + RMANOO sur VTL Quantum DXi directement depuis le serveur d'Aubervilliers (ou Vélizy en cas de bascule site) qui héberge la base.

Ainsi, le BCD n'est plus utilisé par la sauvegarde ; il garde cependant l'avantage d'un retour arrière à J - 1 plus rapidement qu'une restauration.

J'ai fourni aux DBA les paramètres à renseigner dans RMANOO afin d'utiliser le Media Manager NBU en place de TiNa :

```
BLKSIZE=1048576
NB_ORA_SERV='head -n 1 /usr/opensv/netbackup/bp.conf | cut -d" " -f3`
NB_ORA_CLIENT='grep CLIENT_NAME /usr/opensv/netbackup/bp.conf | cut -d" " -f3`
NB_ORA_POLICY=<nom de la police NetBackup à fournir par DEI>
NB_ORA_SCHED=<nom du schedule NetBackup à fournir par DEI>
NB_ORA_TRACE=0
```

Ceci a permis de libérer :

- l'espace utilisé par la plupart des BCV sur le stockage SAN (certains ont été conservés sur les bases les plus sensibles afin de permettre une restauration à J-1).
- les serveurs «Infocentre» (un par OS) qui étaient utilisés sur le site de secours pour monter la base le temps de la sauvegarde.

Pour conclure, les modes opératoires de sauvegarde écartés sont :

- en production : le RMAN Server Less,
- en PPMCE : l'agent TiNa for Oracle,
- en PPHOM : la sauvegarde à froid (base arrêtée),

sur TiNa et STK (robots physiques) ;

et les modes opératoires de sauvegarde retenus pour être généralisés sont :

- en production et en PPMCE : le RMAN LAN Free,
- en PPHOM : la sauvegarde à froid (base arrêtée),

sur NBU et DXi (VTL).

Au vu de sa volumétrie, la base OTAP DMC sera sauvegardée via le réseau IP.

b) DB2

Symantec propose à son catalogue un agent NBU pour DB2. Celui-ci n'ayant jamais été utilisé chez France Telecom, il doit faire l'objet d'une qualification.


J'ai délégué cette opération à une autre personne de l'équipe, qui a procédé aux tests de qualification et fourni les documents d'installation et d'exploitation, que j'ai vérifié et validé.


c) MySQL


Symantec ne proposant pas d'agent NBU pour MySQL, la solution retenue sera d'embarquer les dumps lors des sauvegardes de fichiers plats (comme sous TiNa).

6.8. Reporting - Supervision des sauvegardes

Les sauvegardes Tina de la plate-forme MDSP  sont actuellement supervisées dans un outil appelé «Troisoo» développé en interne.

Les sauvegardes Tina et NBU des plates-formes de services des IAS  sont actuellement supervisées dans un outil appelé «Infsvg» développé en interne.

J'ai présenté l'outil «Infsvg» aux acteurs du projet qui ne le connaissaient pas, et formé les exploitants MDSP  à sa consultation.

Au fur et à mesure de la migration, il me faudra vérifier la remontée des informations des clients NBU MDSP  dans l'outil «Infsvg».

En parallèle du projet de migration, mon équipe a conduit un projet nommé «Protons» de remise à plat de l'outil de supervision des sauvegardes. J'ai ainsi collaboré à la rédaction du cahier des charges.

J'ai également défini avec les exploitants sauvegarde les indicateurs à remonter dans la Boucle Qualité Nationale (BQN) :

- Indicateur de Qualité de service : Taux d'échec = nombre de jobs de sauvegarde en réussite / nombre total de jobs de sauvegarde.
- Indicateur de Performance : «Top 10» des clients présentant les plus mauvais débits de sauvegarde.
- Indicateur de Disponibilité de service en HO et en HNO : DSSMT = taux de Disponibilité de Service par Mois de Toute l'infrastructure de sauvegarde (en %) : DSSMS = taux de Disponibilité de Service par Mois et par Site de l'infrastructure de sauvegarde (en %).
- Indicateur de Traitement des tickets : Nombre de tickets SWAN reçus, nombre de tickets SWAN traités, durée moyenne de traitement.
- Indicateurs de Volumétrie : Volumétrie sauvegardée pour chaque application.
- Nombre de clients déclarés et nombre de clients sauvegardés pour chaque application.

6.9. Découpage du projet en phases - mise au point des procédures générales

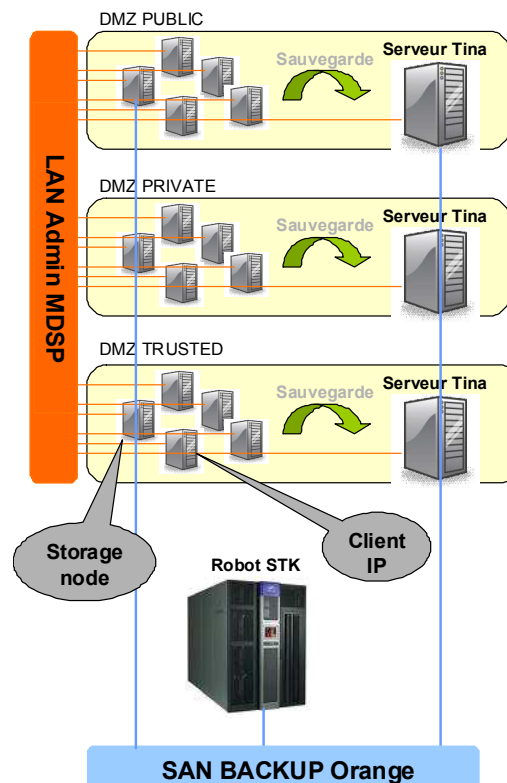


Illustration 38 : phase 0 : état des lieux

Phase 0 : État des lieux

Les clients IP sont sauvegardés via un agent TiNa sur le serveur TiNa présent dans leur DMZ qui envoie les données sur la robotique STK.

Les storage nodes sont sauvegardés via un agent TiNa sur la robotique STK.

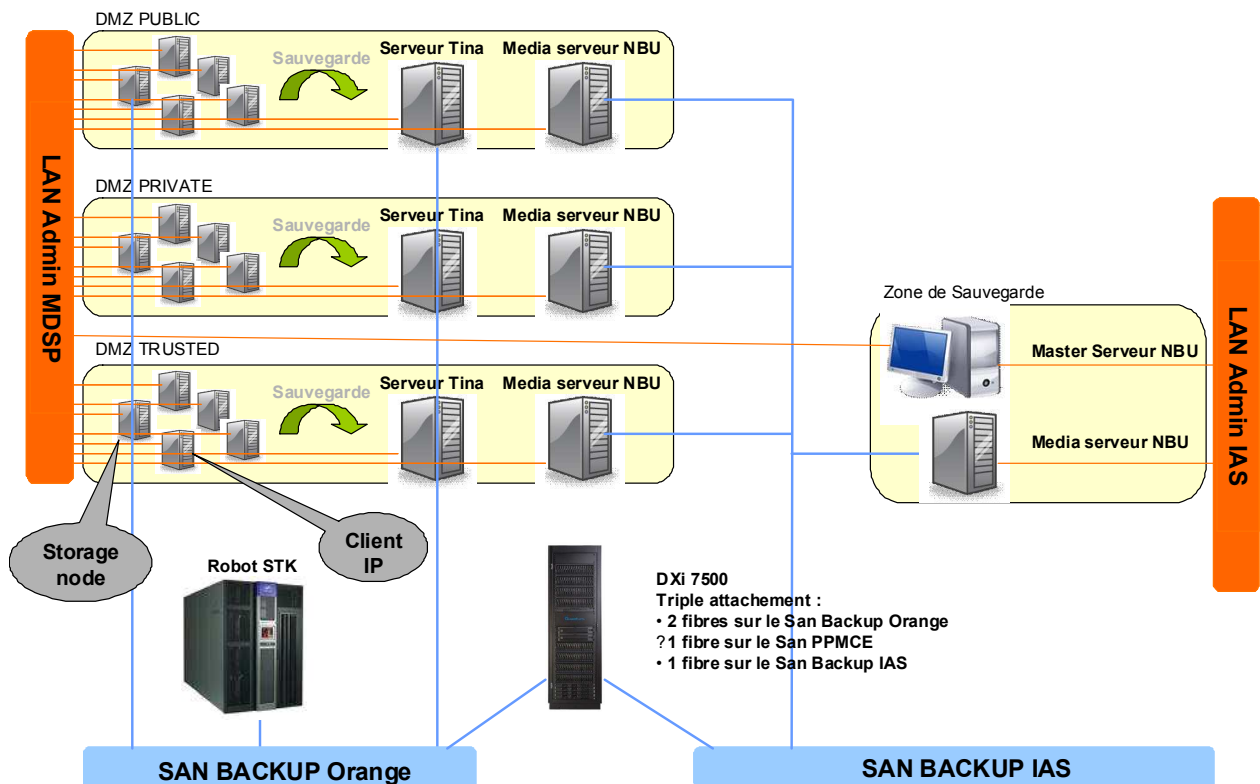


Illustration 39 : phase 1 : déploiement

Phase 1 : Déploiement

On déploie l'infrastructure NBU :

- master serveur dans la ZSV de l'IAS
- 1 media serveur dans l'IAS
- 1 media serveur dans chaque DMZ MDSP
- 1 DXi 7500 rattaché aux différents SAN de sauvegarde

Pas de modification dans la sauvegarde des serveurs MDSP.

En avance de phase, on déploie les agents NBU sachant que :

- un client IP TiNa devient un client IP NBU
- un storage node TiNa devient un SAN media serveur NBU

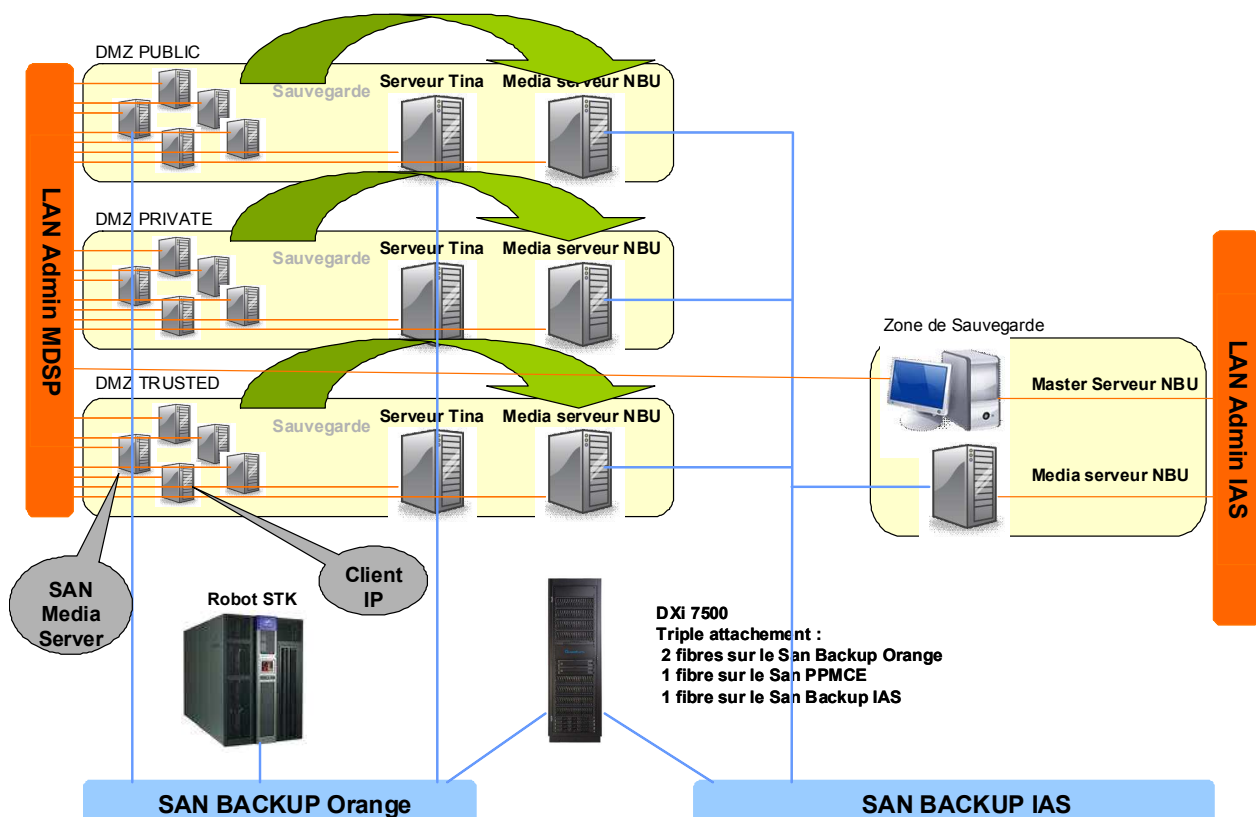


Illustration 40 : phase 2 : migration

Phase 2 : Migration

Les clients IP sont sauvegardés via un agent NBU sur le media serveur (d'infrastructure) NBU présent dans leur DMZ qui envoie les données sur la robotique virtuelle DXi. Les SAN media serveurs sont sauvegardés via un agent NBU sur la robotique virtuelle DXi. Dans les deux cas, les méta-données sont envoyées au master.

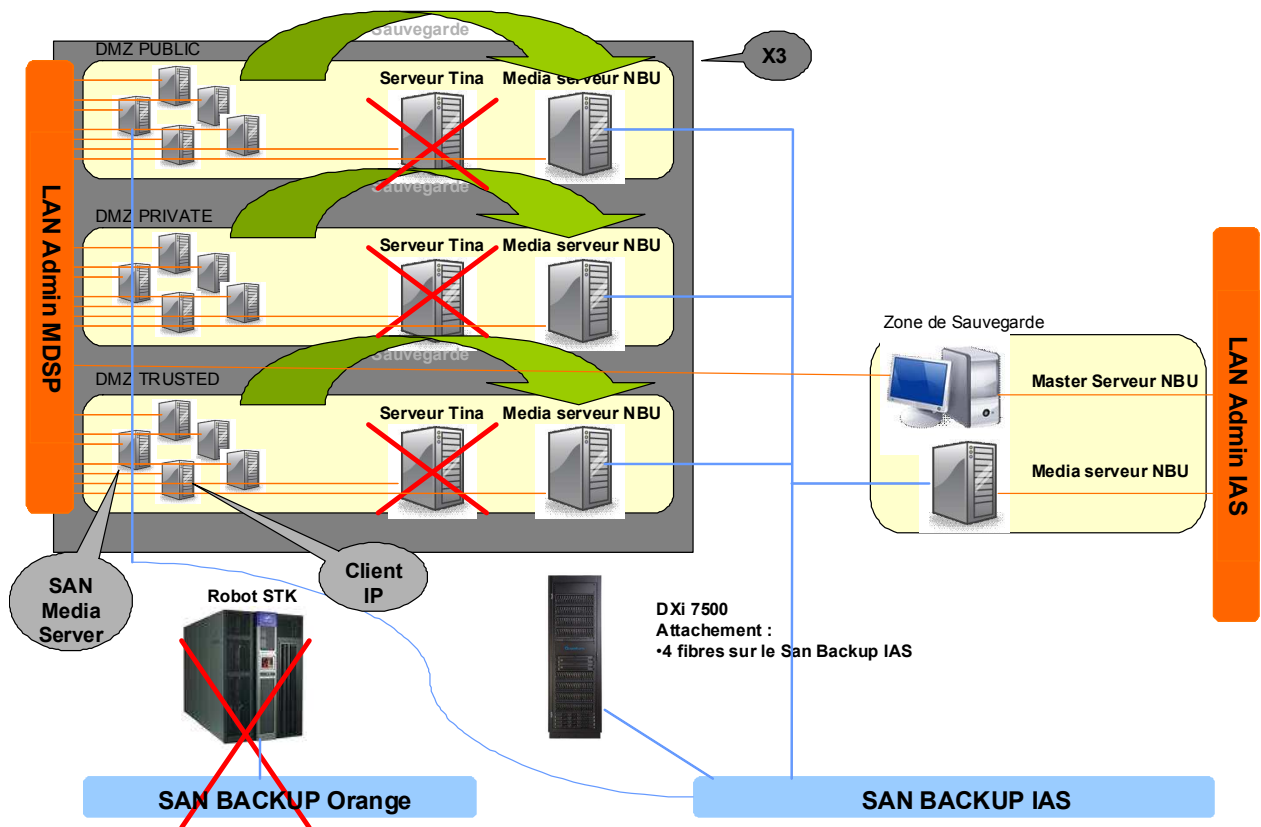


Illustration 41 : phase 3 : fin du projet

Phase 3 : Fin du projet

En fin de rétention des sauvegardes TiNa :

- on supprime les agents TiNa et leurs pré-requis (filesystem, flux...)
- on déplace les fibres des SAN medias serveurs depuis le SAN Backup Orange vers le SAN Backup IAS
- on déplace les fibres du DXi depuis le SAN Backup Orange vers le SAN Backup IAS (la fibre du DXi connectée au SAN PPMCE restant en place)
- on libère les serveurs TiNa, la robotique STK et le SAN Backup Orange

7. MISE EN ŒUVRE DE LA SOLUTION

7.1. Planning

Dates clés	Jalons projet TTM (cf. chapitre 4.1)
5 Février 2008	T0 Lancement du projet (accord formel MOA sur le périmètre : cadre fonctionnel général, coûts, délais)
15 Mai 2008	T1 Revue de conception - Accord pour démarrer la phase de réalisation
30 Septembre 2008	T2 Revue de développement - Accord pour démarrer la phase de recette
9 mars 2009	T3 Revue de lancement sur le marché ou généralisation - Accord pour ouverture généralisée du service
9 avril 2009	T4 Fin du projet, fin de la généralisation et activités transférées en activités récurrentes

Tableau XV : macro-calendrier / jalons exprimés dans le PMP

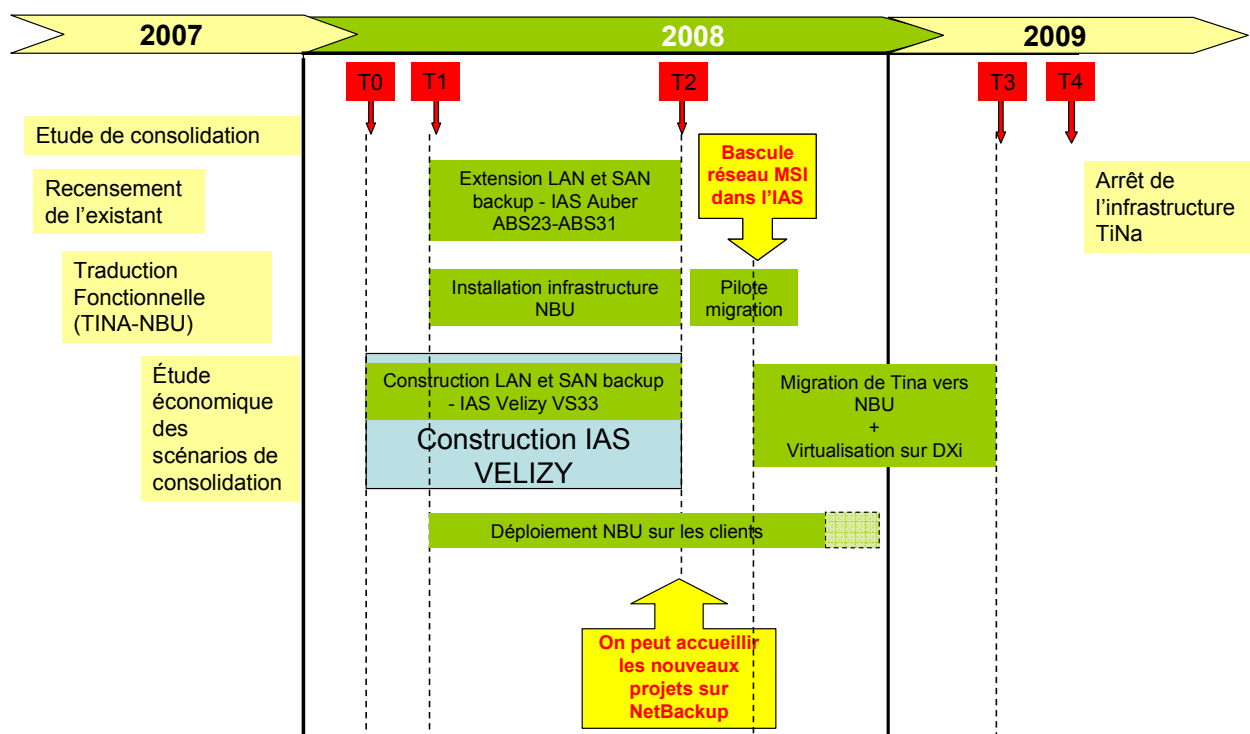


Illustration 42 : roadmap

Je vais reprendre dans les chapitres suivants (7.2 à 7.6) le découpage du projet en phases décrit dans le chapitre 6.9.

Phase 1 : Déploiement

- **Commande puis installation des 2 DXi 7500**
 - sur Aubervilliers : DXi 7500 24TB Brut / 18TB Utile 40 VTD
 - sur Vélizy : DXi 7500 24TB Brut / 18TB Utile 120 VTD
- **Extension réseau IAS** vers les salles MDSP
- **Déploiement de l'infrastructure NBU** :
 - Sur Aubervilliers :
 - 3 medias serveurs dédiés MDSP : 1 dans chaque DMZ
 - 1 DXi 7500 : dans l'IAS
 - Sur Vélizy :
 - 1 master mutualisé MDSP (Production, Pré-production Homologation, Pré-production Maintenance) et tous projets IAS : dans la ZSV IAS
 - 9 medias serveurs dédiés MDSP : 1 dans chaque DMZ de chacune des 3 plates-formes
 - 1 media serveur mutualisé IAS : dans la ZSV IAS
 - 1 DXi 7500 : dans l'IAS
 - Interconnexion des SAN de sauvegarde MDSP et IAS
 - Transcription des stratégies de sauvegarde TiNa en polices NBU
 - Configuration infra NBU (partitionnement DXi, polices, zoning, VTD, STU...)
- **Déploiement sur les clients** :
 - Installation de l'agent NBU et de l'Uproc \$U pour NBU sur les clients (pour rappel : un client IP TiNa devient un client IP NBU, alors qu'un storage node TiNa devient un SAN media serveur NBU)
 - Zoning et déclaration des VTD sur les storage nodes TiNa (qui deviennent SAN medias serveurs NBU)
 - Déploiement de RMAN sur les clients avec l'agent TiNa for Oracle

Phase 2 : Migration

- **Pré-requis** :
 - Communication (routage et flux) :
 - sauvegarde TiNa : internes au nid, à conserver pour le retour arrière et les restaurations sur TiNa
 - sauvegarde NBU : internes au nid, et entre les nids et la ZSV
 - exploitation TiNa : STP doit pouvoir continuer à exploiter TiNa dans l'IAS
 - réplication DXi : sur la Passerelle Inter VPN (PIV) entre les 2 IAS (ports TCP et UDP 58345)

- **Actions de migration :**
 - Test des flux de sauvegarde :
 - TiNa : pour le retour arrière et les restaurations sur TiNa -> interne nid MDSP
 - NBU : interne nid MDSP, et entre le nid MDSP et la ZSV
 - Mise à jour des fichiers hosts
 - Test de sauvegarde et de restauration NBU
 - Activation des polices NBU et désactivation des stratégies TiNa
 - Mise à jour de l'ordonnancement des Uprocs \$U sur les clients
- **Actions post-migration :**
 - Vérification du bon fonctionnement des sauvegardes NBU
 - Après la fin des rétentions TiNa, suppression de l'agent TiNa, J4 des serveurs TiNa et J4 des lecteurs physiques

7.2. Déploiement de l'infrastructure de sauvegarde

J'ai validé avec les fournisseurs Quantum et IBM les configurations exactes à commander en fonction des besoins exprimés au chapitre 6.3 et du **catalogue top-sourcing** :



- 1 DXi 7500 24TB Brut / 18TB Utile 40 VTD
- 1 DXi 7500 24TB Brut / 18TB Utile 120 VTD
- 1 IBM x3950-M2 / 2 Xeon (quadricore) 1,6Ghz / 4 Go de RAM
- 9 IBM x3950-M2 / 2 Xeon (quadricore) 1,6Ghz / 8 Go de RAM / 2 QLogic 4Gb FC
- 9 baies IBM EXP-3000 SAS avec 12 disques de 300 Go à 15000 tours
- 9 cartes IBM ServeRAID MR10M SATA-300 / SAS - PCI Express x8




J'ai présenté en CI2A le projet, décrit le matériel que je souhaitais implanter et en particulier **l'espace et la consommation électrique** qu'il allait occuper et consommer. Dans l'IAS de Vélizy, le DXi 7500 a été installé dans son rack (livré par le fournisseur Quantum) ; une contrainte interne au Datacenter d'Aubervilliers nous a obligé à faire racker le DXi 7500 par Quantum dans une baie Atos 47U, ce qui nous a permis d'y ajouter d'autres éléments, mais a également nécessité de s'assurer du maintien des conditions de garantie et de maintenance auprès de Quantum. Les autres éléments (serveurs et baies) sont implantés dans des racks existants.




Équipement	Hauteur (U)	Conso (kW)	Nombre d'équipements	Total hauteur (U)	Total conso (kW)
DXi 7500	36	2,76	2	72	5,52
EXP3000	2	0,5	9	18	4,5
x3950-M2	4	2,88	10	40	28,8
PW450	4	1,1	4	16	4,4
				146	43,22

Tableau XVI : espace et consommation électrique des éléments de l'infrastructure NBU

J'ai déposé les hostnames des serveurs (cf. tableaux XVII à XIX) en m'appuyant sur les conventions de nommage :

- pour la plate-forme MDSP  : **m**sp **a** **b** **x** **yy** avec
 - a** = Site
 - a = Aubervilliers
 - v = Vélizy
 - b** = environnement
 - i = INT
 - r = PPERF et PPMCE
 - q = PPHOM
 - p = PROD
 - x** = zone
 - x = 3 => zone Public
 - x = 5 => zone Private
 - x = 7 => zone Trusted
 - x = 8 => zone Publique-RSC
 - x = 9 => zone Administration
 - yy** = numéro incrémental
- pour les IAS  : **o**sias **a** **yy** avec
 - a** = Site
 - a = Aubervilliers
 - v = Vélizy
 - yy** = numéro incrémental

Le site de Vélizy étant un nouveau site IAS , tous les hébergements ont été réalisés dans les salles voisines VS32 (salle historique MDSP ) et VS33 (nouvelle salle IAS ). Sur ce site, nous n'avons pas rencontré d'autre problème que d'attendre l'aménagement de la salle VS33 (implantation des baies, alimentation électrique, câblages...).

Le site d'Aubervilliers étant déjà un site IAS , les hébergements ont été réalisés dans les salles distantes ABS11 et ABS23 (salles historiques IAS ) et ABS31 (salle historique MDSP ) qui a été étendue. Sur ce site, nous avons dû attendre l'extension de la salle ABS31 mais aussi gérer les problèmes de raccordement inter-salle via des rocares.

Afin de suivre le déploiement de l'infrastructure de sauvegarde sur les 2 sites, j'ai mis au point le tableau de suivi suivant :

Fonction	Environnement	DMZ	Site	Type	Hostname	SN	Salle cible	Baie cible	Dalle	Position U	Providence OS	Providence câblage
Media	PPHOM	Public	Vélizy	X3950-M2	mspvq347	99B0048	VS32	V32	C9L28	1	213183	207721
Baie disques	PPHOM	Public	Vélizy	EXP3000		130542L	VS32	V32	C9L28	5		
Media	PPHOM	Private	Vélizy	X3950-M2	mspvq636	99B0047	VS32	V32	C9L28	7	213184	207721
Baie disques	PPHOM	Private	Vélizy	EXP3000		130374N	VS32	V32	C9L28	11		
Media	PPHOM	Trusted	Vélizy	PP450	mspvq760	6B03142982	VS31	V52	C16L45	5	213190	207721
Media	PPMCE	Public	Vélizy	X3950-M2	msprv364	99B0043	VS32	V42	C18L24	5	213185	207721
Baie disques	PPMCE	Public	Vélizy	EXP3000		130402R	VS32	V42	C18L24	9		
Media	PPMCE	Private	Vélizy	X3950-M2	msprv636	99B0026	VS32	V47	C14L49	2	213186	207721
Baie disques	PPMCE	Private	Vélizy	EXP3000		130402Y	VS32	V47	C14L49	16		
Media	PPMCE	Trusted	Vélizy	PP450	msprv779	6B03142920	VS32	V40	C16L24	9	213191	207721
Media	PROD	Public	Vélizy	X3950-M2	mspv389	99B0042	VS32	V28	C23L24	14	213187	207721
Baie disques	PROD	Public	Vélizy	EXP3000		130402W	VS32	V28	C23L24	20		
Media	PROD	Private	Vélizy	X3950-M2	mspv670	99B0039	VS32	V28	C23L24	10	213188	207721
Baie disques	PROD	Private	Vélizy	EXP3000		130374P	VS32	V28	C23L24	18		
Media	PROD	Trusted	Vélizy	PP450	mspv787	6BR3517506	VS32	V15	C32L28	5	213192	207721
Media	PROD	Public	Aubervilliers	X3950-M2	mspap389	99B0027	ABS31	A33	AQ63	23	218701	210724
Baie disques	PROD	Public	Aubervilliers	EXP3000		130386Y	ABS31	A33	AQ63	27		
Media	PROD	Private	Aubervilliers	X3950-M2	mspap666	99B0037	ABS31	A33	AQ63	31	218699	210724
Baie disques	PROD	Private	Aubervilliers	EXP3000		130385D	ABS31	A33	AQ63	29		
Media	PROD	Trusted	Aubervilliers	PP450	mspap785	6BR3517507	ABS31	A33	AQ63	16	214581	210724
Master	PROD	IAS	Vélizy	X3950	osiasv01	99T2993	VS33		L49C23	17	platonisation ok	214416
Media	PROD	IAS	Vélizy	X3950	osiasv10	99T3177	VS33		L49C26	17	platonisation ok	
Baie disques	PROD	IAS	Vélizy	EXP3000		130600A	VS33		L49C26	21		
VTL	PROD	IAS	Aubervilliers	DXi 5500	osiasa06		ABS31		AK56		ok 6/6/8	
VTL	PROD	IAS	Aubervilliers	DXi 7500	osiasa07	CX0807BVA00017	ABS31		AK56		ok	
VTL	PROD	IAS	Velizy	DXi 7500	osiasv02	CX0809BVA00027	VS33		L45C25		ok 6/6/8	
Master	PROD	IAS	Auber		osiasa01				AM40		ok	214418

Tableau XVII : implantation physique de l'infrastructure NBU

Je pouvais ainsi identifier rapidement un équipement, et suivre les demandes de racking, de câblage ou d'installation de l'OS.

La plate-forme MDSP ne proposant pas d'accès à distance, seuls les masters et le media serveur dédié pour l'IAS de Vélizy devaient bénéficier d'une carte RSA connectée au VLAN RSA IAS. En tant qu'ancien exploitant, je n'étais que trop conscient de la nécessité d'avoir une prise en main à distance sur tous les serveurs de l'infrastructure de sauvegarde NBU. J'ai ainsi étudié avec les architectes réseaux la possibilité de connecter les cartes RSA des serveurs IBM et les cartes XSCF des serveurs Fujitsu sur le VLAN IAS. Ceci a dû faire l'objet d'une **dérogation** de sécurité, car cela permet l'accès à la plate-forme MDSP sans passer par le bastion MDSP. J'ai dû **justifier** que l'accès est encore plus restreint car il nécessite une **authentification via TDIMG** avant une authentification sur le serveur sur lequel on veut prendre la main.

Ces cartes d'accès à distance se sont d'ailleurs révélées très utiles lors du déploiement, en particulier lors de la mise à jour des firmwares et de la configuration du bios des cartes FC, et ce sans solliciter inutilement le niveau 0 (les équipes de proximité, cf. le chapitre 8.3).

J'ai demandé les adresses IP suivantes :

- une adresse sur le réseau d'administration MDSP (de chaque DMZ) pour les medias serveurs MDSP,
- une adresse IP pour la prise en main à distance (RSA sur les serveurs IBM, XSCF sur les serveurs Fujitsu) dans le VLAN RSA IAS,
- une adresse dans le VLAN d'administration sauvegarde IAS pour le master de l'IAS de Vélizy,
- une adresse dans le VLAN service sauvegarde IAS pour le master de l'IAS de Vélizy,
- une adresse dans le VLAN d'administration sauvegarde IAS pour chaque DXi,
- une adresse dans le VLAN service sauvegarde IAS pour chaque DXi (pour de futurs accès NAS que je n'utiliserai cependant pas dans le cadre de ce projet).

Les VLAN 204 (administration sauvegarde) et 209 (service sauvegarde) de l'IAS d'Aubervilliers étant pleins, j'ai demandé leur extension aux architectes réseaux. Le VLAN 204 (administration sauvegarde) a ainsi été étendu avec le VLAN 3604, et le VLAN 209 (service sauvegarde) a été étendu avec 3605.

Fonction	Environnement	DMZ	Site	Type	Hostname	@ip Admin	VLAN admin	@ip Service	VLAN service	@ip RSA ou XSCF
Media	PPHOM	Public	Vélizy	X3950-M2	mshpvq347	10.162.12.70 / 26	969			10.97.161.43
Baie disques	PPHOM	Public	Vélizy	EXP3000						
Media	PPHOM	Private	Vélizy	X3950-M2	mshpvq636	10.162.12.200 / 25	970			10.97.161.44
Baie disques	PPHOM	Private	Vélizy	EXP3000						
Media	PPHOM	Trusted	Vélizy	PP450	mshpvq760	10.162.13.41 / 26	971			10.97.161.45
Media	PPMCE	Public	Vélizy	X3950-M2	mshpr364	10.162.252.69 / 25	625			10.97.161.46
Baie disques	PPMCE	Public	Vélizy	EXP3000						
Media	PPMCE	Private	Vélizy	X3950-M2	mshpr636	10.162.253.138 / 24	626			10.97.161.47
Baie disques	PPMCE	Private	Vélizy	EXP3000						
Media	PPMCE	Trusted	Vélizy	PP450	mshpr779	10.162.252.210 / 25	627			10.97.161.48
Media	PROD	Public	Vélizy	X3950-M2	mshvp389	10.162.8.98 / 25	965			10.97.161.49
Baie disques	PROD	Public	Vélizy	EXP3000						
Media	PROD	Private	Vélizy	X3950-M2	mshvp670	10.162.9.172 / 24	966			10.97.161.50
Baie disques	PROD	Private	Vélizy	EXP3000						
Media	PROD	Trusted	Vélizy	PP450	mshvp787	10.162.8.205 / 25	967			10.97.161.51
Media	PROD	Public	Aubervilliers	X3950-M2	mshpap389	10.162.6.104 / 25	961			10.93.36.183
Baie disques	PROD	Public	Aubervilliers	EXP3000						
Media	PROD	Private	Aubervilliers	X3950-M2	mshpap666	10.162.7.172 / 24	962			10.93.36.184
Baie disques	PROD	Private	Aubervilliers	EXP3000						
Media	PROD	Trusted	Aubervilliers	PP450	mshpap785	10.162.140.8 / 25	890			10.96.36.185
Master	PROD	ZSV IAS	Vélizy	X3950-M2	osiasv01	10.97.160.115/28 LAN admin global MDSP :	1409	10.97.160.179/28	1701	10.97.160.163/28
Media	PROD	ZSV IAS	Vélizy	X3950-M2	osiasv10	10.97.160.114/28	1409	10.97.160.178/28	1701	10.97.160.162/28
Baie disques	PROD	ZSV IAS	Vélizy	EXP3000						
VTL	PROD	ZSV IAS	Aubervilliers	DXi 5500	osiasa06	10.97.156.98/28	3604	10.97.156.113/28	3605	
VTL	PROD	ZSV IAS	Aubervilliers	DXi 7500	osiasa07	10.97.156.97/28	3604	10.97.156.114/28	3605	
VTL	PROD	ZSV IAS	Vélizy	DXi 7500	osiasv02	10.97.160.116/28	1409	10.97.160.180/28	1701	
Master	PROD	ZSV IAS	Auber		osiasa01	10.96.32.52/29 LAN admin global MDSP :	204	10.96.90.17/28	209	10.96.36.176

Tableau XVIII : adressage IP de l'infrastructure NBU

J'ai sous-traité aux Pilote de Mise En Production (PMEP) de chaque site, la production des documents suivants :

- schémas d'implantation dans les baies
- schémas de câblage Ethernet et SAN

J'ai fourni les instructions afin d'installer la baie de disques EXP3000 contenant 12 disques de 300 Go (15000 tours) soit une volumétrie brute de 3,6 To pour avoir une volumétrie utile de 2,8 To avec 1 disque de parité et 1 disque hot-spare :

- récupérer l'utilitaire StorMan sur le site d'Adaptec (www.adaptec.com puis sélectionner les rubriques Support, Downloads, Serial ATA II Raid, Unified Serial SAS / SATA, Adaptec RAID 4800 SAS, Storage Manager Downloads)
- installer l'utilitaire StorMan : # rpm -Uvh asm_linux_x86_v5_20_17414.rpm
- créer le groupe Raid avec 1 disque de parité et 1 disque hot-spare via la commande : # /usr/StorMan/arconf CREATE 2 LOGICALDRIVE Stripesize 256 Name Staging Method Build Rcache RON Wcache WBB MAX 5EE 0,8 0,9 0,10 0,11 0,12 0,13 0,14 0,15 0,16 0,17 0,18 0,19
- créer les file systems pour héberger le disk staging via les commandes : #pvcreate /dev/sdb ; vgcreate stagingvg /dev/sdb ; lvcreate -L 700G -n staging10s stagingvg ; lvcreate -L 700G -n staging16j stagingvg ; mke2fs /dev/stagingvg/staging10s ; mke2fs /dev/stagingvg/staging16j ; mkdir -p /data/nbu/staging10s ; mkdir -p /data/nbu/staging16j ; mount /dev/stagingvg/staging10s /data/nbu/staging10s ; mount /dev/stagingvg/staging16j /data/nbu/staging16j
- éditer le fichier /etc/fstab pour que les fs soient montés automatiquement au reboot.

J'ai supervisé l'installation de l'OS Platon Linux palier G7 (Red Hat Advanced Server 4) et Platon Solaris palier G5 (Solaris 9). Comme présenté au chapitre 6.3.c, j'ai demandé à mettre en place l'agrégation de liens sur les medias serveurs Linux. Le mode retenu, 4 ou 802.3ad, offre la répartition de charge sur 2 liens avec doublement de la bande passante. J'ai demandé aux architectes réseaux la documentation pour mettre en place le trunking sur les switches ; j'ai fourni les modifications à apporter au fichier /etc/modprobe.conf sur les serveurs :

- alias bond0 bonding
- options bond0 miimon=100 mode=4 primary=eth4
- install bond0 /sbin/modprobe -a eth4 eth5 && /sbin/modprobe bonding

J'ai supervisé l'installation du PLI NetBackup Master et Media Serveur 6MP4. En plus du fichier de suivi présenté ci-dessus, j'ai rapidement dû créer un nouveau fichier d'avancement afin d'alerter sur les blocages qui risquaient de mettre en péril le planning du projet et de l'ouverture de l'IAS de Vélizy. De plus, le suivi fut complexifié par le fait que les demandes de travaux diffèrent entre les IAS et la plate-forme MDSP.

Equipement	Caracteristiques	Tache 0 : RACKAGE + Energie		Tache 1: Installation OS		Tache 2: Cablage Ethernet		Tache 3: ouverture flux +		Tache 4: Netbackup licences		Tache 5: Netbackup Installation		Tache 6: SAN cablage	
		Porteurs: D.Boucard / T.Trebuchet		Porteur: STP - Y.Ouled Amor/Mourad DJOUAD		Porteur: IAS ITE - D.Boucard/T.Trebuchet		Porteur: DISU - P.Dounga		Porteur: ITE F.Matzinger		Porteur: TMC Infra - B.Saulme/C.Fossier/C.Delhomme		Porteur: IAS IEP - D.Boucard	
		Demande	Retour	Demande	Retour	Demande	Retour	Demande	Retour	Demande	Retour	Demande	Retour	Demande	Retour
osiasv01	x3950 Master Velizy Linux	15/04/2008	12/06/2008	13/06/2008	17/06/2008	06/06/2008	20/06/2008	20/06/2008	02/07/2008	09/06/2008	11/06/2008	02/07/2008	18/07/2008	28/04/2008	12/06/2008
osiasv10	x3950 Media Velizy Linux	15/04/2008	12/06/2008	13/06/2008	20/06/2008	06/06/2008	20/06/2008	20/06/2008	02/07/2008	09/06/2008	11/06/2008	02/07/2008	18/07/2008	28/04/2008	12/06/2008
osiasv02	Dxi 7500 Velizy	15/04/2008	15/04/2008	Quantum	GA le 06/06	06/06/2008	20/06/2008	20/06/2008	02/07/2008	09/06/2008	11/06/2008	02/07/2008	18/07/2008	28/04/2008	12/06/2008
mspv6389	x3950 Media Velizy Linux	15/04/2008	18/07/2008	Providence 213187	livraison + rackage x3950 le 11/07/2008	15/04/2008	30/04/2008	17/06/2008		09/06/2008	11/06/2008			28/04/2008	12/06/2008
mspv670	x3950 Media Velizy Linux	15/04/2008	18/07/2008	Providence 213188	livraison + rackage x3950 le 11/07/2008	15/04/2008	30/04/2008	17/06/2008		09/06/2008	11/06/2008			28/04/2008	12/06/2008
mspv787	pp450 Media Velizy Solaris	15/04/2008	19/05/2008	26/05/2008	17/06/2008	15/04/2008	30/04/2008	17/06/2008		09/06/2008	11/06/2008			28/04/2008	12/06/2008
mspv364	x3950 Media Velizy Linux	15/04/2008	18/07/2008	Providence 213185	livraison + rackage x3950 le 11/07/2008	15/04/2008	30/04/2008	17/06/2008		09/06/2008	11/06/2008			28/04/2008	12/06/2008
mspv636	x3950 Media Velizy Linux	15/04/2008	18/07/2008	Providence 213186	livraison + rackage x3950 le 11/07/2008	15/04/2008	30/04/2008	17/06/2008		09/06/2008	11/06/2008			28/04/2008	12/06/2008
mspv779	pp450 Media Velizy Solaris	15/04/2008	19/05/2008	26/05/2008	17/06/2008	15/04/2008	30/04/2008	17/06/2008	en cours	09/06/2008	11/06/2008			28/04/2008	12/06/2008
mspv347	x3950 Media Velizy Linux	15/04/2008	18/07/2008	Providence 213183	livraison + rackage x3950 le 11/07/2008	15/04/2008	30/04/2008	17/06/2008		09/06/2008	11/06/2008			28/04/2008	12/06/2008
mspv636	x3950 Media Velizy Linux	15/04/2008	18/07/2008	Providence 213184	livraison + rackage x3950 le 11/07/2008	15/04/2008	30/04/2008	17/06/2008		09/06/2008	11/06/2008			28/04/2008	12/06/2008
mspv760	pp450 Media Velizy Solaris	15/04/2008	19/05/2008	26/05/2008	17/06/2008	15/04/2008	30/04/2008	17/06/2008		09/06/2008	11/06/2008			28/04/2008	12/06/2008
oiasv1a	Switch SAN Backup Velizy	15/04/2008	02/05/2008			15/04/2008	27/06/2008	17/06/2008	20/06/2008					28/04/2008	20/06/2008
mspap389	x3950 Media Auber Linux	15/04/2008	13/06/2008	Providence à faire	livraison + rackage x3950 le 11/07/2008	17/06/2008	25/06/2008			09/06/2008	11/06/2008			28/04/2008	17/06/2008
mspap666	x3950 Media Auber Linux	15/04/2008	13/06/2008	Providence à faire	livraison + rackage x3950 le 11/07/2008	17/06/2008	25/06/2008			09/06/2008	11/06/2008			28/04/2008	17/06/2008
mspap785	pp450 Media Auber Solaris	15/04/2008	01/06/2008	10/06/2008	en attente Interconnexion ABR31 - ABR23 (*) et (**)	17/06/2008	25/06/2008			09/06/2008	11/06/2008			28/04/2008	17/06/2008
osiasa06	Dxi 7500 Auber	15/04/2008	15/04/2008	Quantum	GA le 06/06	15/04/2008	30/04/2008			09/06/2008	11/06/2008			28/04/2008	17/06/2008
oiasa1b	Switch SAN Backup Auber	15/04/2008	11/06/2008			15/04/2008	en cours							28/04/2008	en cours

Tableau XIX : suivi du déploiement de l'infrastructure NBU

En tant que MOE📖, j'ai été personnellement en charge de la commande et de la gestion des licences pour l'infrastructure de sauvegarde :

- 4 «NetBackup Server, UNIX, Enterprise Server» pour les 4 medias serveurs📖 Solaris,
- 10 «NetBackup Server, Linux, Enterprise Server» pour le master📖 serveur de Vélizy et les 9 medias serveurs Linux,
- 2 «NetBackup Option, Cross-Platform, Virtual Tape Option, 1 TB» pour les DXi📖 ; j'ai en effet convenu avec l'éditeur Symantec d'initier une licence minimale pour chaque DXi📖 puis de régulariser en fin de projet. Ce taux de remplissage des DXi a pu être facilement suivi grâce à l'outil DxiStor développé en interne par un de mes collègues (*cf. chapitre 8.5*).

J'ai en outre dû veiller à ce que les serveurs migrés soient bien munis des licences nécessaires :

- les clés de licences pour les clients IP «simples» et les agents de base de données (Oracle📖 et DB2📖) étant installées sur le master📖 serveur, rien n'empêche techniquement de sauvegarder ces machines alors que la licence n'a pas été acquise auprès de l'éditeur Symantec,
- les clés de licences pour les SAN media serveurs📖 sont installées localement ; des retards de commande des licences NBU ont nécessité de recourir à des licences temporaires, et donc de suivre leur remplacement par des clés définitives ou leur prolongation à l'aide d'une nouvelle clé temporaire.

J'ai également été en charge des demandes de supervision de l'infrastructure de sauvegarde auprès des 2 équipes de supervision IAS📖 et MDSP📖. En plus de la supervision «standard» des OS, j'ai défini les objets à superviser spécifiquement pour le master📖 NetBackup :

- les daemons pbx_exchange (communication EMM), nbemm (base EMM), bprd (requêtes), bpdbm (catalogue) et vmd (base des volumes),
- les file systems /exec/products/netbackup, /exec/products/netbackup/logs et /data/NBK1 (catalogue).

et pour les medias serveurs NetBackup :

- les daemons pbx_exchange (communication EMM), ltid (bras robotique), vmd (base des volumes) et avrd (lecture des codes-barres des bandes),
- les file systems /exec/products/netbackup, /exec/products/netbackup/logs, /data/nbu/staging10s et /data/nbu/staging16j.

Comme exposé au chapitre 6.1, j'ai été en charge de toutes les demandes d'ouverture de flux pour le fonctionnement de NetBackup auprès des 2 équipes réseaux IAS📖 et MDSP📖 ; j'ai en ce sens rédigé les Demandes d'Ouverture de Flux (DOF) pour les exploitants réseau MDSP📖 et les matrices de flux pour les exploitants réseau IAS📖 en spécifiant toutes les sources et destinations :

- pour les flux TCP 22 (ssh), TCP 13722 (bpjava), TCP 25 (smtp) entre les bastions et les serveurs d'infrastructure NBU
- pour les flux TCP 1556 (pbx_exchange) et TCP 13724 (vnetd) entre les serveurs d'infrastructure NBU et les autres serveurs de la plate-forme MDSP📖
- pour les flux TCP 80 et 58345 entre les 2 DXi📖
- pour le flux TCP 22 (ssh) entre le master📖 NBU de Vélizy et le serveur de reporting «Infsvg»

Comme exposé au chapitre 6.3.a, j'ai fourni les paramètres pour les DXi, à savoir une partition VTL par plate-forme, configurée avec 3 VTD par media serveur d'infrastructure et 2 VTD par SAN media serveur (sauf pour la base SwapCom à laquelle on affectera 3 VTD par nœud, conformément au nombre de channels utilisés actuellement pour sauvegarder cette bases de données sous Tina). Le bras de chaque VTL sera piloté par le media serveur de la zone privée de la plate-forme associée, afin de limiter le nombre de partitions, tout en séparant les données de production de celles de pré-production.

On aura donc sur osias04 (le DXi 7500 d'Aubervilliers) :

- 1 VTL pour la plate-forme de production MDSP avec 9 VTD

Et sur osiasv02 (le DXi 7500 de Vélizy) :

- 1 VTL pour la plate-forme de production MDSP avec 9 VTD
- 1 VTL pour la plate-forme PPMCE MDSP avec 9 VTD
- 1 VTL pour la plate-forme PPHOME MDSP avec 9 VTD
- 1 VTL pour la plate-forme de production IAS avec 3 VTD

7.3. Pilote

Le but du pilote consiste à vérifier :

- le bon fonctionnement des modes opératoires de déploiement et d'activation par les exploitants,
- le bon fonctionnement de la chaîne de sauvegarde NBU via des tests de sauvegarde et de restauration,
- l'absence de dégradation de performances par rapport aux sauvegardes Tina.

J'ai défini l'infrastructure nécessaire pour monter le pilote :

- un master serveur,
- un media serveur,
- une librairie DXi.

Pour des raisons de délai, nous avons utilisé l'infrastructure de l'IAS d'Aubervilliers, déjà opérationnelle avec un DXi 5500.

Le pilote a été réalisé sur quelques machines et a été concluant.

7.4. Principe d'installation des packages clients et migration des clients

a) Pré-requis

- Vérifier que le client fait partie du périmètre
- Vérifier le bon fonctionnement des sauvegardes Tina
- Vérifier la version de NBU à installer
- Vérifier l'espace disque disponible
- Créer le filesystem /exec/products/netbackup
- Mettre à jour le fichier hosts
- Configuration réseau
- Vérifier les flux

b) Installation

- Transférer le package client (récupéré sur Ogre)
- Installer et configurer le PLI📖 (selon la documentation Platon📖)
- Vérifier la configuration
- Tester la communication avec le master📖
- Tester la communication avec le media serveur📖 associé
- Faire un test de sauvegarde
- Faire un test de restauration
- Configuration des polices NBU selon l'EB
- Mettre en place les exclusions
- Mettre en place les pré- et post-traitement

c) Activation - post installation

- Désactiver les stratégies Tina
- Activer les polices NBU
- Vérifier le bon fonctionnement des sauvegardes NBU le lendemain
- Désinstaller l'agent Tina à la fin des rétentions
- Supprimer le système du catalogue Tina à la fin des rétentions

7.5. Principe d'installation des packages SAN media serveurs et migration des storage nodes

a) Pré-requis

- Vérifier que le client fait partie du périmètre
- Vérifier le bon fonctionnement des sauvegardes Tina
- Vérifier la volumétrie sauvegardée (supérieure à 50 Go)
- Vérifier la version de NBU à installer
- Vérifier l'espace disque disponible
- Créer le filesystem /exec/products/netbackup
- Mettre à jour les fichiers hosts
- Configuration réseau
- Vérifier les flux
- Zoner le serveur avec le DXi
- Configurer les lecteurs au niveau de l'OS (*nécessite de redémarrer les serveurs Linux et Solaris*)
- Configurer le persistent binding

b) Installation

- Transférer le package media serveur (récupéré sur Ogre)
- Installer et configurer le PLI (selon la documentation Platon)
- Entrer la clé de licence Media
- Vérifier la configuration
- Tester la communication avec le master
- Configurer les lecteurs au niveau de NBU
- Faire un test de sauvegarde
- Faire un test de restauration
- Configuration des polices NBU selon l'EB
- Mettre en place les exclusions
- Mettre en place les pré- et post-traitement

c) Activation - post installation

- Désactiver les stratégies Tina
- Activer les polices NBU
- Vérifier le bon fonctionnement des sauvegardes NBU le lendemain
- Désinstaller l'agent Tina à la fin des rétentions
- Supprimer le système du catalogue Tina à la fin des rétentions
- Supprimer la zone entre le storage node et la robotique STK
- Faire la demande de déplacement de fibre
- Déplacement fibre du SAN MDSP vers SAN IAS

7.6. Principe d'installation des agents de base de données et migration des sauvegardes des bases de données

- Vérifier que le client fait partie du périmètre
- Vérifier le bon fonctionnement des sauvegardes RMAN ou DB2 Tina
- Transférer le package Oracle ou DB2 pour NBU (récupéré sur Ogre)
Installer et configurer l'agent Oracle ou DB2 pour NBU (selon la documentation Platon)
- Faire un test de sauvegarde
- Désactiver les stratégies Tina
- Activer les polices NBU
- Configurer les Uprocs \$U pour NBU

Voir le mode opératoire détaillé en annexe 5.

7.7. Problèmes rencontrés

J'ai déjà exposé au chapitre 7.2 certains problèmes rencontrés lors du déploiement, dont certains avant été heureusement identifiés en amont du projet (*cf. chapitre 5.6*) :

- le retard dans la disponibilité des salles,
- le blocage par le CI2A de l'installation du DXi 7500 dans le Datacenter d'Aubervilliers,
- la mise en place d'un accès à distance pour les serveurs d'infrastructure NBU,
- l'interconnexion des SAN de sauvegarde.

Certains problèmes identifiés en amont du projet (*cf. chapitre 5.6*) ne se sont finalement pas confirmés :

- le DXi 7500 a été livré en version bêta. Aucune garantie n'était alors proposée par Quantum quant aux données que nous pourrions y stocker. La General Availability (GA) a finalement été livrée avant le début de la migration.
- l'exploitation du serveur Tina dans les IAS n'a pas généré de problème d'accès, étant donné que le projet de migration réseau a été abandonné (*cf. chapitre 6.4*).

J'ai également dû faire face aux problèmes suivants apparus en cours de projet :

a) Changements d'équipe

Des réorganisations d'équipes au sein du groupe France Telecom ont eu lieu en cours de projet. Par exemple :

- les exploitants NBU de TMC/INFRA ont été transférés chez DEI/SEO,
- l'activité des DBA a été transférée en Roumanie...

Par ailleurs, certains externes, dont le chef de projet, n'ont pas vu leur contrat reconduit jusqu'au terme du projet.

Ces changements de direction et/ou d'interlocuteurs ont pénalisé le projet, car il a fallu redéfinir les actions et les responsabilités.

b) Xinetd

Les exploitants m'ont remonté un problème de communication entre l'infrastructure de sauvegarde NBU et certains serveurs Linux. Après analyse, toute connexion sur le port TCP 13724 correspondant au service vnetd était rejetée sur ces serveurs. Ceci était dû à une configuration particulièrement restrictive du daemon xinetd lequel gère la connexion au processus vnetd. En accord avec les experts sécurité, nous avons validé la modification des fichiers /etc/xinetd.conf et /etc/xinetd.d/[bv]* comme suit :

- ajouter les noms des services «vnetd, bpcd, vopied et bpjava-msvc» à la ligne «enabled =>» dans /etc/xinetd.conf
- commenter les lignes max_load et per_source dans l'/etc/xinetd.conf
- ajouter la ligne «onlyfrom = <mastername> <medianames> <clientname_a>» à /etc/xinetd.d/vnetd
- ajouter la ligne «onlyfrom = <clientname_a>» à /etc/xinetd.d/bpcd
- ajouter la ligne «onlyfrom = <clientname_a>» à /etc/xinetd.d/vopied
- ajouter la ligne «onlyfrom = <clientname_a>» à /etc/xinetd.d/bpjava-msc

Voir l'annexe 6 pour plus d'informations sur l'xinetd.

c) Retard dans la rédaction des EB

Les EB📖 devaient être initialement rédigées par les MOE📖 applicatives, lesquelles connaissent les fichiers à restaurer en cas de problème sur leurs serveurs, et leurs besoins en termes de fréquence et rétention. Cette tâche a malheureusement été négligée et un externe a été recruté en cours de projet pour y pallier. J'ai assuré le transfert de compétences afin qu'il puisse transposer les EB📖 Tina en EB📖 NBU. Les fréquences et rétentions standards des IAS📖 ont été imposées, sauf justification motivée. Nous avons ainsi dû proposer de l'archivage à certains projets qui nécessitaient des rétentions longues.

d) Interactions avec d'autres projets

Comme exposé au chapitre 2.3, le projet Backup interagit avec d'autres projets du programme MIO, ce qui a amené divers problèmes :

- **network migration to IAS** : j'ai déjà exposé au chapitre 6.4 la révision de l'architecture réseau cible suite au NOGO de migration des infrastructures réseaux IP et SAN📖 de MDSP📖 vers IAS📖, ce qui m'a amené à émettre une dérogation.
- **virtualisation** : les serveurs qui ont été virtualisés bénéficient d'une sauvegarde système via le VCB📖 qui peut être redondante avec les sauvegardes en place sous Tina ; il a donc fallu en tenir compte dans la reprise des EB📖. Ce travail a été rendu délicat par le fait de changement du statut d'éligibilité de certains serveurs à la virtualisation.
- **decommissionning** : les serveurs qui devaient être décommissionnés étaient initialement exclus du projet Backup. Les modifications apportées au périmètre de ce projet (retard dans

l'arrêt des services, démarrage de nouvelles applications...) ont directement impacté le périmètre du projet Backup.

- **AIX consolidation** : des storage nodes Tina sur serveurs physiques AIX ont été virtualisés sur des partitions de PL1660, sans leur laisser de carte HBA. Or si le PL1660 dispose de cartes HBA que l'on peut partager via un VIOs, celles-ci ne permettent que de présenter des LUN de type disque (de baie SAN), et non des périphériques de type lecteurs de bande. Des incidents de sauvegarde sont donc survenus sur les premiers serveurs concernés, aussi bien sur les storage nodes Tina que sur les SAN media serveurs NBU. La sauvegarde des partitions via IP étant exclue (au vu des volumétries et performances attendues), la fonctionnalité de NPIV n'étant pas encore disponible (et de surcroît non qualifiée) et aucune qualification de sauvegardes de partitions AIX n'ayant été émise, nous avons dû ajouter des cartes HBA sur les PL1660 dont une a été dédiée par storage node ou SAN media serveur. Depuis, suite à une demande d'Ecocenter, j'ai qualifié la sauvegarde de partitions AIX via un media serveur NBU installé dans une partition du châssis, les données à sauvegarder empruntant exclusivement le réseau interne du châssis.

e) Problème de restauration des bases de données Oracle en cluster

Le produit RMAN étant complexe, les DBA utilisent souvent des scripts pour le manipuler. Son packaging officiel chez France Telecom se nomme RMAN Lite, avec lequel l'interfaçage avec les Media Managers NBU et Tina est validé par Platon, et les MOE Oracle et Sauvegarde ; j'ai moi-même déjà acquis une bonne connaissance de l'interfaçage de ces 2 produits, grâce à mon expérience au sein du groupe France Telecom. Les DBA MDSP utilisaient cependant un packaging de scripts nommé RMANOO lequel fonctionnait également à priori correctement avec les Media Managers NBU et Tina.

J'avais demandé en début du projet la mise en conformité avec les préconisations groupe, c'est-à-dire l'abandon de RMANOO au profit de RMAN Lite. Suite aux changements d'équipe (cf. chapitre 7.7.a), cette migration n'a pas été effectuée et il est apparu que si RMANOO fonctionnait correctement avec le Media Manager NBU pour des bases Oracle en standalone, ce n'était pas le cas des bases Oracle en cluster. En effet, il n'était pas possible de restaurer les archives logs du 2^{ème} nœud lorsque l'on lançait une restauration de la base depuis le 1^{er} nœud. En collaboration avec les DBA, nous avons trouvé une solution palliative qui consiste à générer le script de restauration sans l'exécuter, puis à l'éditer en renseignant le nom du 2^{ème} nœud sur au moins un channel.

8. FIN DU PROJET

8.1. Bilan de projet

a) Architecture mise en place

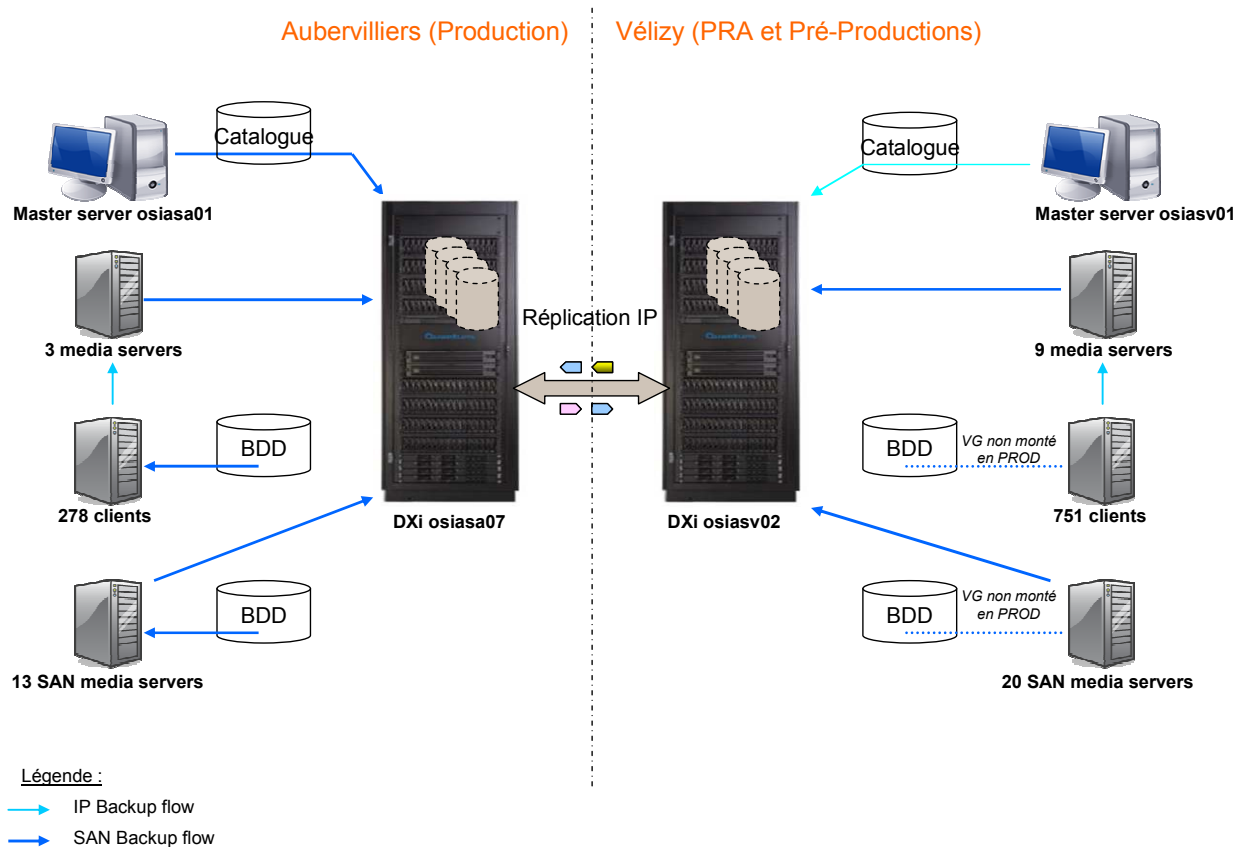


Illustration 43 : architecture mise en place

b) Synthèse globale

Item	😊 ou 😞	Fait notable Bonne pratique ou difficulté Actions mises en oeuvre et leur efficacité	Axes d'amélioration Capitalisation possible
Migration de Tina vers Netbackup	😊	Tous les serveurs pérennes ont été migrés de l'outil de sauvegarde Tina vers l'outil de sauvegarde (groupe) Netbackup.	▶ Les migrations des clusters de bases Oracle ont nécessité davantage d'effort
Sauvegardes sans BCV	😊	▶ la suppression du BCV pour certaines sauvegardes a nécessité : <ul style="list-style-type: none"> ▶ Une modification de l'ordonnement ▶ La mise en place de réplication intersite des sauvegardes ▶ La fiabilisation du lien SRDF 	▶ Un plan d'action a permis d'assurer la suppression du BCV
Dépendances avec les autres projets : AIX et Décommissionnement	😊	▶ Prise en compte des serveurs en cours de décommissionnement ▶ Ajout de cartes HBA (prérequis du projet backup) pris en charge par le projet AIX	▶ La coordination entre projet a pu se faire grâce à la structure en programme (programme MIO)
Démontage de l'infra Tina et désinstallation des agents	😞	Certains serveurs (CSPDB, Mail/PIM et Webmethods) sont toujours sauvegardés sur Tina jusqu'à leur décommissionnement Juin 2009 : l'infra Tina sera démontée hors projet	▶ Pris en charge par la maintenance évolutive
Impacts sur process et organisation des exploitants	😞	Ces impacts n'ont pas été pris en compte dans le projet	▶ l'aspect « Change Management » de l'exploitant ne doit pas être laissé de côté

Tableau XX : synthèse globale

Écart délai : +15% (15 mois au lieu de 13)

Écart coût : +1% (*chiffres initial et final confidentiels*)

c) Gestion des risques

Risque identifié (cause et impact)	Action prévue Mise en oeuvre	Comité d'où est issu l'action	Faits notables ayant fait évoluer le risque	Conclusions (efficacité des actions, points à améliorer,...)
Risque de non disponibilité de DEPFS et DEI	▶ Le planning a tenu compte de la charge acceptée par l'exploitant : ~25 Client IP + ~6 SAN Media par semaine	Réunion hebdomadaire		Nécessité de tenir un planning détaillé
Adhérences avec le projet AIX	▶ Gestion au niveau programme	Réunion d'équipe	Ajout de cartes HBA supplémentaires sur les PL1660 de Velizy et Auber	Des arbitrages ont été nécessaires au niveau programme

Tableau XXI : gestion des risques

d) Bilan pilotage projet



Jalons	Date cible	Date réelle	Ecart	Fait notable ou difficulté justifiant l'écart Actions mises en œuvre, leur efficacité
T0	5-févr.-08	5-févr.-08	0 jours	
T1	15-mai-08	15-mai-08	0 jours	
T2	30-sept.-08	22-oct.-08	+ 1 mois	<ul style="list-style-type: none"> - Le retard de l'installation de l'infra à Vélizy n'a pas permis de passer le pilote avant semaine 31. - L'abandon du projet de migration réseau dans l'IAS en semaine 39 a nécessité une nouvelle architecture soumise à dérogation.
T3	9-mars-09	30-avril-09	+2 mois	<ul style="list-style-type: none"> - La migration des SAN media serveurs a démarré en semaine 3. - Installation de cartes HBA supplémentaires sur PL1660 en semaine 9. - Ajout de serveurs additionnels OWL Core, OWL GW, HHD...

Tableau XXII : bilan pilotage projet

8.2. Gains

État des lieux TiNa + STK	Cible NBU + DXi	Gains attendus
19 catalogues TiNa	2 catalogues NBU	Exploitation simplifiée
Robotiques (STK9310 et SL8500) et bandes physiques	Robotiques (2 DXi 7500) et bandes virtuelles	Fiabilisation (moins de pannes matérielles)
22 lecteurs STK 9940B partagés	Jusqu'à 2 x 160 lecteurs virtuels dédiés	Meilleure parallélisation
Infrastructure et exploitation dédiée MDSP	Infrastructure et exploitation mutualisée IAS	Mutualisation des ressources
~160 k€ / an de maintenance logicielle	0 k€ / an de maintenance logicielle (jusqu'en 2010)	Diminution des coûts d'OPEX
Sauvegarde «en Y» via le SAN	Réplication via IP	Diminution des coûts réseau
Compression	Déduplication	Diminution de la volumétrie stockée.
Sauvegarde des bases Oracle via BCV ou agent spécifique	Sauvegarde des bases Oracle via RMAN Lan Free	Standardisation des mécanismes de sauvegarde

Tableau XXIII : gains

Le DXi  offre de plus la possibilité de faire du NAS  c'est-à-dire de faire des écritures directement en NFS/CIFS sans passer par un logiciel de sauvegarde (ex : export de dump de sauvegarde système).



L'arrêt des sauvegardes Tina a permis le décommissionnement de l'infrastructure de sauvegarde Tina :


- Décommissionnement des 11 partitions/serveurs Tina et «Infocentre»,
- Libération du stockage occupé par les serveurs Tina pour héberger les catalogues et l'espace cache utilisé par le macro-multiplexage, et réattribution à d'autres applications,
- Décommissionnement des switches Brocade 12000 et 24000,
- Réattribution des 22 lecteurs de bande STK à d'autres projets Orange (par ex. Content Billing).

Type machine Puissance (en watts) Action	Demandeur	Acteur	Avancement par serveur									
			PP250N	PP450	x345	x336	PL3200N	PL3200N	PL3200N	PL3200N	PL3200N	PL3200N
			800	1100	350	600	5000	5000	5000	5000	5000	5000
attente fin de rétention			mshpr752	mshpr738	mshpr740	mshpr758	mshpr774	mshpr760	mshpr774	mshpr730	mshpr748	mshpr738
arrêt RCAT			NA	NA	NA	NA	ok	ok	NA	ok	NA	ok
arrêt supervision	Florent	500038	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
arrêt svgs (y compris croisées) + arrêt daemon	Florent	DEI014	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
dezonning serveurs	Florent	DEI004	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
libération lecteurs	Florent	DEI014	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
libération stockage	Florent	DEI004	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
chgt mdp root + arrêt OS	Florent	DEI014	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
décâblage Ethernet + alimentation	Thierry		ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
dezonning lecteurs + décâblage Fibre	Manuel	DEI004	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
CI2A	Thierry		ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
libération ressources (hostnames, IPs)	Thierry		ok	ok	ok	ok	ok	ok	ok	ok	ok	ok
fermeture des flux	Florent	DEI020	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok

Tableau XXIV : décommissionnement des serveurs Tina

L'optimisation de la sauvegarde des bases de données a permis en particulier de libérer :

- l'espace utilisé par la plupart des BCD  sur le stockage SAN  (certains ont été conservés sur les bases les plus sensibles afin de permettre une restauration à J-1).
- les serveurs «Infocentre» (un par OS) qui étaient utilisés sur le site de secours pour monter la base le temps de la sauvegarde.

Cette solution a également permis de **réduire les coûts d'exploitation** car elle est depuis longtemps la **référence au sein du SI  de France Telecom** et les exploitants y sont fortement rodés.

8.3. Mise en place de la gestion récurrente

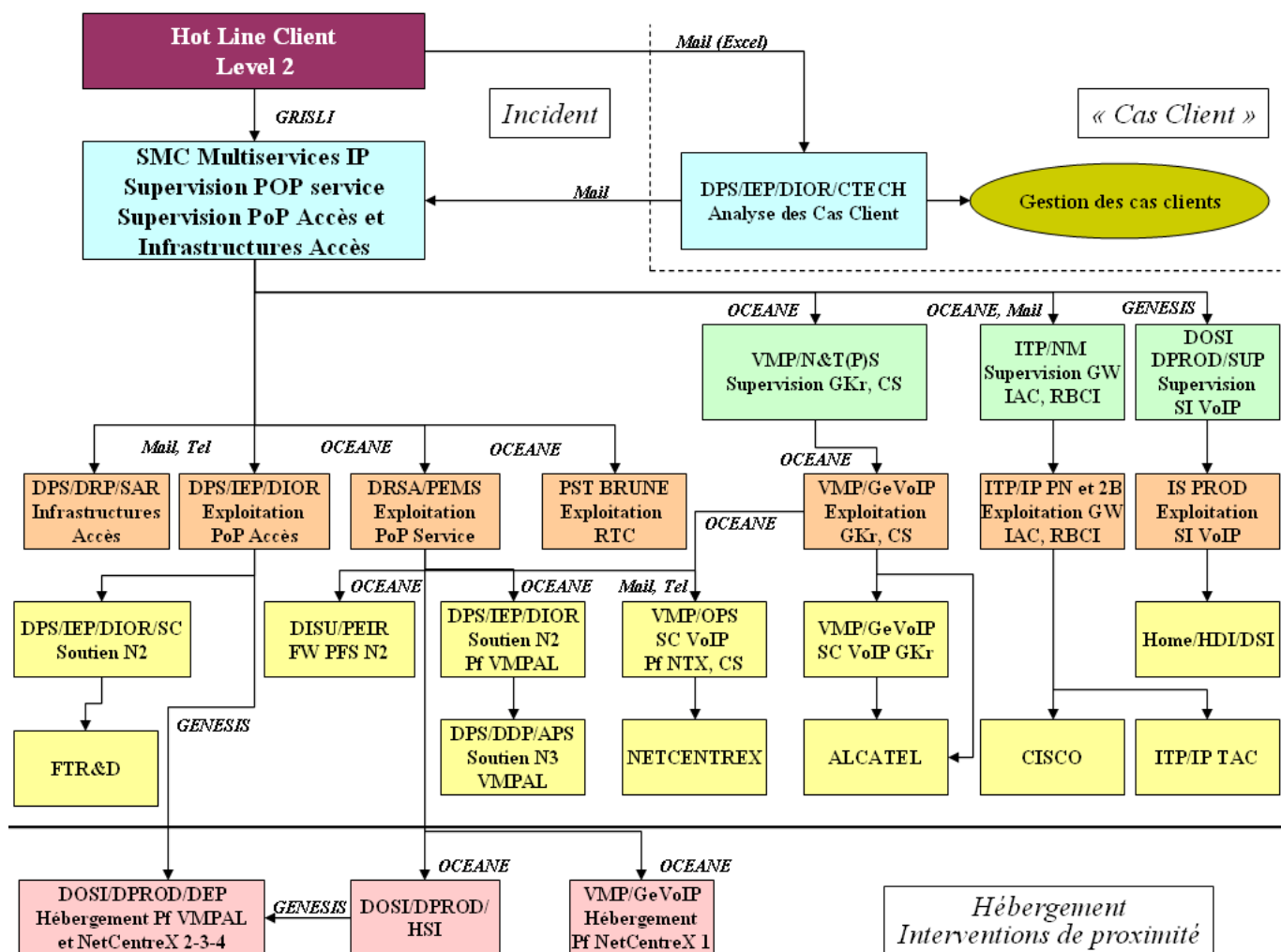


Illustration 44 : définition de la chaîne de soutien

Présentation de l'organisation par niveaux :


Niveau 0 : La proximité correspond à la fourniture des services de prestation d'intervention physique sur les infrastructures matérielles des plates-formes. La proximité assure aussi l'ensemble des prestations d'hébergement des plates-formes de service ; elle est notamment garante de la Qualité de Service des infrastructures du site d'hébergement



Niveau 1 : Il correspond au premier niveau de compétence demandé pour la maintenance du bon fonctionnement et correspond pour partie aux tâches d'exploitation courantes menées par un technicien d'exploitation.

Le niveau 1 a la charge de surveiller les services et systèmes de la plate-forme, de réaliser des contrôles et des travaux planifiés et d'agir lors d'un incident ou d'une demande d'intervention sur une application client. Les modes opératoires sont alors documentés dans le(s) dossier(s) d'Exploitation : fiches consignes, documents solutions, fiches d'exploitation. Il pilote les équipes d'intervention de proximité.

Le N1 est généralement l'exploitant des applications et infrastructures. Il est garant de la Qualité de fonctionnement des équipements des plates-formes et applicatifs qu'il exploite. À ce titre, il doit maintenir une visibilité ainsi qu'une capacité au diagnostic sur l'ensemble des composants des plates-formes de service.

Niveau 2 : Il nécessite une connaissance globale des architectures déployées sur les différents sites de production, ainsi qu'une compétence technique pointue sur les logiciels et matériels des plates-formes.


Il assure la gestion technique du périmètre, en s'appuyant d'une part sur les centres de support fournisseurs et d'autre part sur les équipes de conception (MOE , architectes).



Le niveau 2 se tourne vers sa MOE  pour les PLI .

Le N2 assure la gestion technique du périmètre. Il est garant de la cohérence d'ensemble des plates-formes et valide l'ensemble des opérations réalisées sur ce parc. Le N2 n'est pas un exploitant, c'est une équipe unique de support technique au service de l'exploitation.

8.4. Transfert de compétence aux équipes France Telecom

J'ai assuré le transfert de compétences aux différentes équipes :

- niveau 1 pour qu'ils soient autonomes pour restaurer leurs données,
- niveau 2, en particulier dans le cadre d'un test de PRA  (restauration de données d'Aubervilliers sur le site de Vélizy).

J'ai par la suite donné une formation à l'outil NetBackup aux équipes SSPO en Roumanie qui ont repris l'exploitation applicative et systèmes des serveurs MDSP , et assurent donc le support de 1^{er} niveau de l'outil NBU  sur ces serveurs.

8.5. Qualité de service - Mise en place de la BQN

À partir de l'outil «Infsvg», on peut extraire les statistiques des jobs de sauvegarde afin d'évaluer la qualité de service et la publier dans la Boucle Qualité Nationale (BQN) :

Mois	PPHOM	PPMCE	PROD
	% sauvegardes réussies	% sauvegardes réussies	% sauvegardes réussies
janv-09	87,6	92,7	96,9
févr-09	93,4	93,8	97,9
mars-09	92,7	96,3	98,7
avr-09	96,3	97,4	99,3
mai-09	98	96,9	98,1
juin-09	91,5	96,8	98
juil-09	82	93,8	97,3
août-09	84,1	94,7	98,8
sept-09	95,7	95,4	98,7
oct-09	96,2	97,7	99,4

Tableau XXV : taux de réussite des jobs de sauvegarde

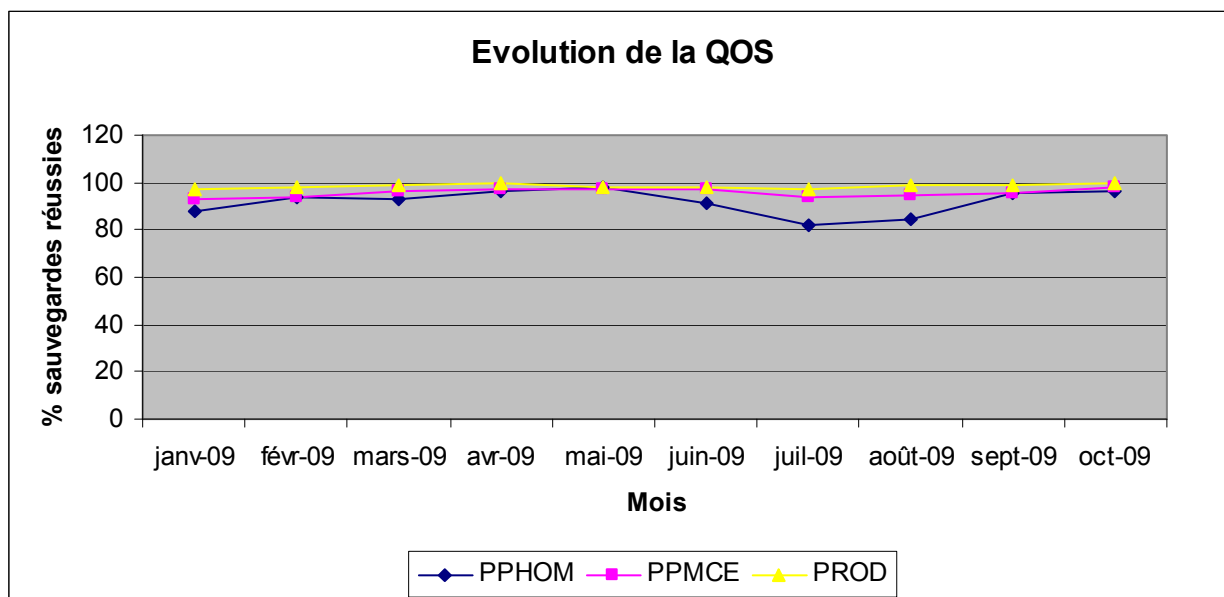


Illustration 45 : qualité de service

À partir de l'outil DXiStor développé par l'équipe ITE/AS, on peut surveiller le taux d'occupation et de déduplication des DXi :



Illustration 46 : suivi des DXis

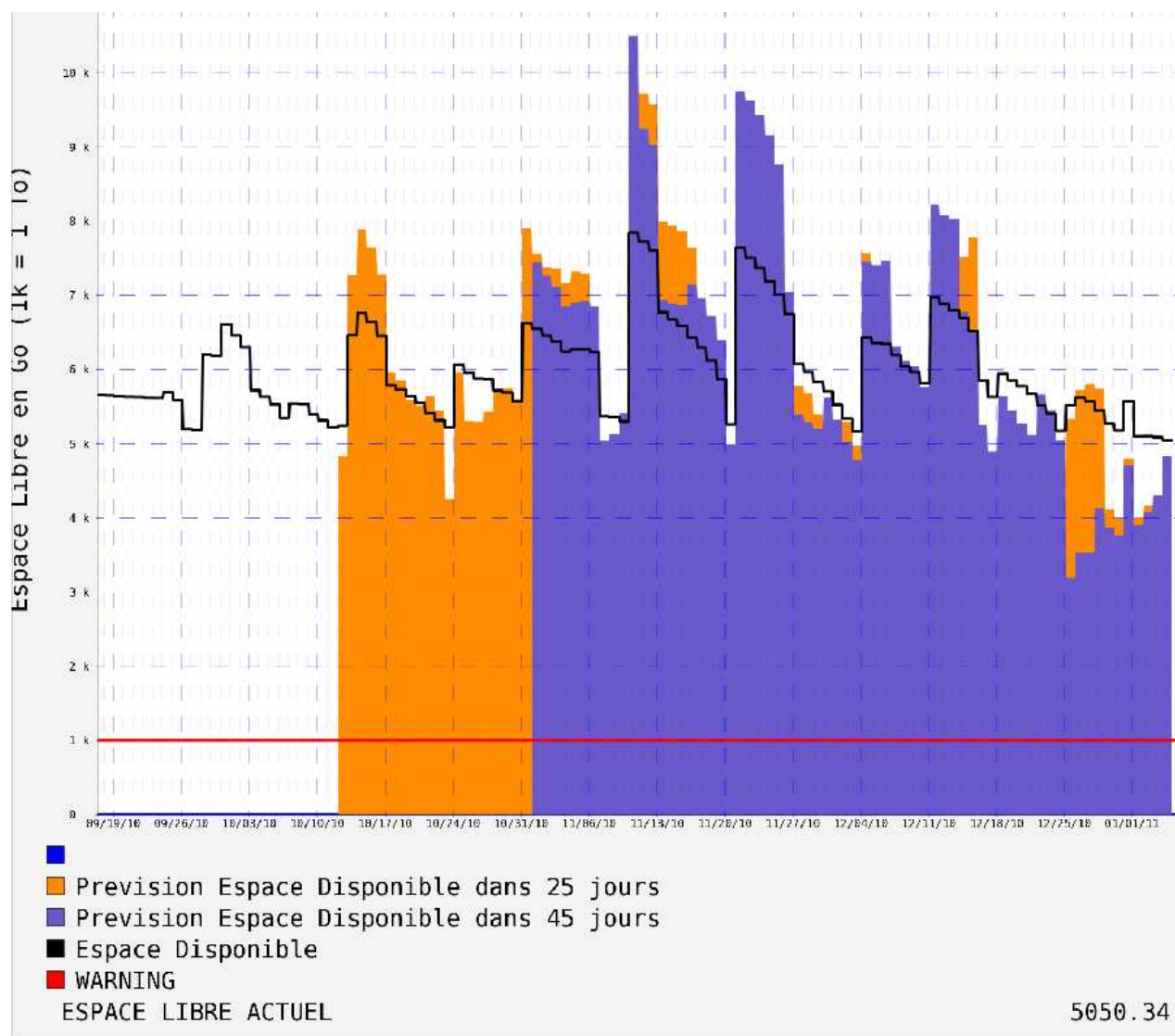


Illustration 47 : projection des consommations d'un DXi

9. CONCLUSION

9.1. Bilan personnel - améliorations

Le recours à une **équipe dédiée**, disposant de tous les accès nécessaires aux infrastructures Tina et NBU et aux clients de la plate-forme MDSP📖, aurait été à mon avis bien plus profitable que d'essayer d'impliquer toutes les différentes équipes d'exploitation, qui ont d'ailleurs été soumises à réorganisation pendant le projet. Cette organisation avait bien fonctionné lors du projet de migration mené à la DOSI📖, même s'il a fallu gérer des modifications de mot de passe (accès temporaires fournis à l'équipe de migration).

J'aurai également souhaité pouvoir **accéder par moi-même** aux plates-formes, surtout pendant la phase d'audit. En effet, je me suis basé sur des documents (inventaire de la plate-forme, DAT Tina, documentation RMANOO...) qui ne reflétaient pas la réalité c'est-à-dire qui se sont révélés partiellement **différents de ce qui était effectivement déployé sur la plate-forme**. Des documents ont dû être mis à jour suite à la découverte de **non-conformités** au fur et à mesure de l'avancement du projet. À titre de comparaison, lors du projet de migration mené à la DOSI📖, j'ai personnellement mis à jour les procédures de migration pendant la phase pilote que j'ai conduite, ce qui a permis de délivrer des documents quasi-définitifs (hormis quelques bugs rencontrés sur des cas spécifiques) à l'équipe de migration.

L'échantillon de serveurs sélectionnés **pour le pilote n'a pas été suffisamment représentatif** à mon avis. En effet, nous n'avions à disposition que 4 clients du même OS et qui n'hébergeaient pas de bases de données. Nous avons ainsi perdu du temps lors de la phase de migration sur **les cas les plus complexes constitués par les SAN media serveurs📖 hébergeant les bases Oracle📖 en cluster📖**.

9.2. Apport du cursus CNAM

Grâce au cursus CNAM, j'ai réussi à **appréhender le processus d'un projet informatique**, à en gérer les différentes étapes et les acteurs, ce qui m'a **permis de gagner en efficacité**. Il m'a permis d'avoir **une meilleure approche et une vision de bout en bout** du déroulement d'un projet informatique, des acteurs et de leur rôle. Cette vision m'a été utile pour **délimiter le périmètre des chantiers à réaliser et les différents acteurs à impliquer et à manager**.

Il m'a **apporté les compétences** nécessaires pour le pilotage d'un projet, (aussi bien dans la rédaction documentaire et les processus qualité que dans le management d'une équipe projet), compétences que j'ai pu mettre en application au cours des différentes contributions réalisées pour ce projet, et ce pour chacune des phases, aussi bien étude que conception / réalisation et mise en production. **La gestion et le suivi des risques** d'un projet au quotidien m'ont été grandement simplifiés grâce aux formations du cursus CNAM.

D'un point de vue personnel, mon entourage a constaté **une amélioration dans ma façon de communiquer** et ma capacité à faire passer un message. J'ai également constaté au cours de ce projet que **les aspects concernant l'exploitabilité des applications ne sont pas suffisamment pris en compte en amont des projets**. Ainsi, les MOE📖 **commencent en général à s'en préoccuper lors de la phase de mise en production, ce qui est trop tardif** : une application dont l'exploitabilité est mal conçue sera délicate à exploiter. En gardant à l'esprit que la durée de vie

d'une application est souvent bien plus importante que le temps consacré à sa conception, on mesure l'impact sur les coûts d'exploitation récurrents d'une mauvaise gestion de l'exploitabilité d'une application. À l'inverse, et à la décharge des MOE📖, **l'exploitant doit de son côté mieux formaliser le processus de mise en production et mieux communiquer auprès des maîtrises d'œuvre** le rôle de chacun des acteurs de l'exploitation.

Ces réflexions m'amènent à conclure **qu'il est important de bien définir les besoins avec la MOA📖 et à impliquer les bons acteurs** afin de permettre une simplification des processus et des interactions entre les différentes entités. **Une forte implication et une grande maîtrise d'un projet de bout en bout** sont d'ailleurs des **enjeux majeurs du cursus CNAM**.

9.3. Remerciements

Je tiens à remercier **l'ensemble des intervenants CNAM** qui, avec grande pédagogie, ont su me transmettre une partie de leur savoir, mon tuteur CNAM, **M. Cédric Kleinpeter** qui m'a soutenu dans la rédaction de ce mémoire, mon maître de stage chez France Telecom, **M. Joseph Rabita** qui m'a permis de réaliser ce projet au sein de son équipe, **l'ensemble des acteurs chez France Telecom** ayant œuvré à ce projet, et enfin **l'ensemble de la promotion CNAM** avec qui j'ai passé des années riches en échange et amitié.

10. BIBLIOGRAPHIE

[1] Jean-Luc Lucas, Une architecture Internet pour le système d'information de France Telecom, Eyrolles, 2001.

[2] Dave High, NetBackup Architecture Overview, Symantec, 2008.

Adresse internet : http://eval.symantec.com/mktginfo/enterprise/white_papers/b-whitepaper_nbu_architecture_overview_12-2008.en-us.pdf ; taille 2,54 Mo.

[3] *Unknown*, NetBackup Backup Planning and Performance Tuning Guide, Symantec, 2007.

Adresse internet : <http://seer.entsupport.symantec.com/docs/281842.htm> ; taille 1,26 Mo.

[4] *Unknown*, NetBackup Compatibility List, Symantec, 2007.

Adresse internet : <http://seer.entsupport.symantec.com/docs/278064.htm> ; taille 0,25 Mo.

Documents Atempo, Symantec et Oracle.

11. GLOSSAIRE

ACSL : Automated Cartridge System Library Software : Serveur pilotant les bibliothèques de marque SUN (anciennement StorageTek ou STK). Il permet de donner une vue logique aux serveurs de sauvegarde à partir de robotiques physiques (exemple : scinder un robot physique en plusieurs partitions afin de le partager entre différents logiciels de sauvegarde).

BCV : Business Continuance Volumes : produit de la société EMC Corporation permettant le clonage d'un espace de stockage.

Capex : Capital Expenditure : budget d'investissement matériel ; correspond aux immobilisations en comptabilité.

CentricStor : Solution de VTL fournie par la société Fujitsu Siemens.

CI2A : Le Comité d'Implantation des Applications et des Architectures valide l'implantation de matériel dans les Datacenters au vu de sa consommation électrique, de sa dissipation thermique (et donc de ses besoins en systèmes de climatisation) et de l'espace occupé. *Se référer au chapitre 4.2b pour une présentation plus approfondie.*

Cluster : En réseau et système, le terme cluster peut désigner une grappe de serveurs (ou ferme de calcul) de deux serveurs au minimum (appelé aussi nœuds) et partageant une baie de disques commune, pour assurer une continuité de service et/ou répartir la charge de calcul et/ou la charge réseau. On parle de cluster actif-actif (tous les nœuds hébergent tout ou partie du service) ou actif-passif (un ou plusieurs nœuds hébergent le service, un ou plusieurs nœuds sont en standby en attente d'une bascule du service).

COGIT : La Cellule Opérationnelle de Gouvernance des Infrastructures Techniques vérifie et valide les besoins d'un projet désirant s'implanter en IAS. *Se référer au chapitre 4.2c pour une présentation plus approfondie.*

CVAT : Comité de Validation des Architectures Techniques : collège d'architectes veillant à la conformité des architectures avec les préconisations groupe. *Se référer au chapitre 4.2a pour une présentation plus approfondie.*

Datacenter : Un centre de calcul et de traitement des données est un service généralement utilisé pour remplir une mission critique relative à l'informatique. Il comprend en général un contrôle sur l'environnement (climatisation, système de prévention contre l'incendie, etc.), une alimentation d'urgence et redondante, ainsi qu'une sécurité physique élevée.

DBA : DataBase Administrator : administrateur de base de données.

DB2 : Système de gestion de bases de données relationnelles d'IBM.

Déduplication : la déduplication (également appelée factorisation ou stockage d'instance unique) est une technique de factorisation des séquences de données identiques afin d'économiser l'espace utilisé. Chaque fichier est découpé en une multitude de tronçons. On associe un identifiant unique à chacun de ces tronçons, les identifiants étant stockés dans un index. La déduplication a pour objectif de ne stocker qu'une seule fois un même tronçon. Aussi, une nouvelle occurrence d'un tronçon déjà présent n'est pas à nouveau sauvegardée, mais remplacée par un pointeur vers l'identifiant correspondant. On utilise généralement la déduplication sur des solutions de type VTL, mais on peut la retrouver également sur du stockage primaire.

Disk Staging : Disque utilisé comme tampon pour les sauvegardes, avant de les déplacer (généralement sur bandes) lors d'une 2^{ème} phase dénommée «destaging».

DMZ : En sécurité informatique, une DeMilitarized Zone est un sous-réseau isolé par un ou des firewall(s) / pare-feu(x).

DOSI : Direction des Opérations du Système d'Informations : Entité en charge du Système d'Information de France Telecom France.

DWDM : Dense Wavelength Division Multiplexing, multiplexeur de longueur d'ondes optique. Permet de faire passer plusieurs liens FC ou Giga Ethernet sur une même fibre optique (en utilisant des lumières de couleurs différentes).

DXi : Solution de VTL fournie par la société Quantum (anciennement Adic).

EB : Une Expression de Besoin de sauvegarde est un document qui décrit pour un projet donné la liste des fichiers à sauvegarder, à quel moment (date, heure, fréquence) et avec quelle rétention (durée de conservation).

Ecocenter : Programme green de France Telecom.

Exchange : Système de messagerie électronique de Microsoft.

Fabric : Ensemble de switches SAN.

FC : Fibre Channel : protocole de communication permettant de relier des systèmes dialoguant via des commandes SCSI.

Firewall (ou pare-feu) : Système interconnectant deux réseaux, comme un routeur, mais avec des fonctions évoluées de filtrage. Toute communication le traversant est contrôlée au moyen de règles définissant ce qui autorisé.

HBA : Host Bus Adapter : carte d'extension qui permet à un serveur d'accéder à un réseau SAN.

IAS : Infrastructure d'Accès Sécurisé : règles d'infrastructure réseau sécurisé défini par France Telecom pour ses plates-formes de service.

iSCSI : Protocole permettant l'encapsulation de commandes SCSI dans le protocole IP.

LAN free : Méthode de sauvegarde/restauration dont le flux transite via le SAN et non via le réseau IP (sauf les méta-données permettant la mise à jour du catalogue).

LUN : Logical Unit Number. Identifiant d'une target SCSI. Une LUN correspond généralement à un volume disque ou à un lecteur de bandes, et est représentée comme un périphérique par l'OS.

Master serveur : Serveur maître de la solution NetBackup, disposant en particulier de la base de données permettant la gestion des sauvegardes et de l'architecture. Cette base s'appelle le catalogue.

MDSP : Mobile Data Service Platform : Plate-forme centralisée intégrant l'ensemble des services mobiles multimédias du Groupe Orange pour ses implantations majeures en Europe. *Se référer au chapitre 2.2 pour une présentation détaillée de la plate-forme.*

Media serveur : Serveur accédant directement à des lecteurs de bandes via le SAN et ayant des droits privilégiés de sauvegarde au niveau NetBackup. Sous NetBackup, on distinguera le SAN Media Serveur (applicatif) pour un serveur à forte volumétrie ou nécessitant un débit de sauvegarde/restauration important et qui se sauvegardera lui-même via le SAN, et le Media Serveur (d'infrastructure) pour un serveur collectant les flux de sauvegarde de clients IP pour les transmettre à la robotique via le SAN.

Méta-données : Informations (nom, chemin, date de sauvegarde, taille...) de chaque fichier sauvegardé, et indexées dans un catalogue de sauvegarde.

MMS : Multimedia Messaging Service : Message incluant un contenu multimédia échangé entre téléphones portables.

MOE : La Maîtrise d'Œuvre est l'expert qui transforme un besoin en une solution SI. Elle propose à la MOA différents scénarios permettant la livraison d'un produit conforme aux exigences de la MOA. De plus, chez FT, la MOE est garante que la solution proposée est mutualisable pour le groupe.

MOA : La Maîtrise d'OuvrAge traduit un besoin métier au travers d'un cahier des charges. Elle demande à la MOE d'assurer la conception, le développement, les tests et la mise à disposition du projet applicatif, avec des exigences de qualité (sécurité, disponibilité...), selon un certain calendrier et dans les limites d'un budget consacré à l'opération. Elle pilote les grandes phases du déroulement du projet, elle réceptionne les résultats intermédiaires ou finaux, et s'assure que les livraisons sont conformes aux engagements.

MySQL : Système de gestion de bases de données relationnelles sous licence GPL (sauf si la base de données est intégrée à un logiciel propriétaire, dans ce cas la licence devient payante).

NDMP : Network Data Management Protocol : Protocole de communication utilisé pour transporter des données entre des filers NAS et des périphériques de sauvegarde.

NetBackup (ou NBU) : Solution industrielle de l'éditeur Symantec permettant de disposer d'un ensemble de fonctionnalités satisfaisant aux besoins de sauvegarde ou de récupération de données dans un environnement hétérogène.

Network Attached Storage (ou NAS) : serveur de fichiers permettant de partager via IP des données entre plusieurs systèmes via des montages NFS ou CIFS.

Opex : Operational Expenditure : budget d'exploitation et récurrent.

Oracle : Système de gestion de bases de données relationnelles de la société du même nom.

OS : Operating System : Système d'exploitation.

Palier : Ensemble cohérent de versions d'outils ou d'applications dont l'interfonctionnement a été validé.

Persistent binding : Association d'une HBA à une LUN. Sans le persistent binding, les descripteurs de périphérique des lecteurs de bande sont créés dans l'ordre de détection des lecteurs de bande ; le persistent binding permet de s'assurer qu'un lecteur de bandes sera toujours identifié par le même descripteur de périphérique.

PFS : Plates-formes de Service : ensemble de serveurs informatiques et de composants techniques associés qui permettent de fournir aux clients finaux les services réseaux et SI.

PIV : La Passerelle Inter VPN permet d'étendre le BEI d'un site IAS à un autre via le VPN PFS, ou d'étendre le BEX d'un site à un autre via le VPN Exploit PFS. *Cf. annexe I*

Platon : Socle technique standardisé et sécurisé permettant d'installer des plates-formes Unix, Windows et Linux sur des serveurs du catalogue top sourcing. Le principe des livrables Platon repose sur trois couches successives, lesquelles doivent être installées dans l'ordre, l'une après l'autre : Noyau, Accueil, PLI.

PLI : Produit Logiciel d'Infrastructure : Tout composant logiciel fournissant un service d'infrastructure (par ex. : sauvegarde, supervision, ordonnancement...).

PRA : Plan de Reprise d'Activité : un plan permettant d'assurer, en cas de crise majeure ou importante d'un centre informatique, la reconstruction de son infrastructure et la remise en route des applications supportées. Il ne faut pas confondre le PRA dont l'objectif est la résilience, avec le Plan de Continuité d'Activité (PCA) dont l'objectif est la robustesse.

RAID : Redundant Arrays of Independent Disks : technologie permettant de stocker des données sur de multiples disques durs afin d'améliorer, en fonction du type de RAID choisi, la tolérance aux pannes et/ou les performances de l'ensemble.

RMAN : Recovery MANager, est un logiciel destiné aux sauvegardes et restaurations des bases de données Oracle. Oracle le fournit depuis la version 8.0, en remplacement de "Enterprise Backup Utility" (en version 7.3.x d'Oracle). RMAN sauvegarde les données (datafiles), les journaux de transactions (archives logs), le fichier de contrôle (control file) et éventuellement le fichier de configuration d'instance (spfile). RMAN peut utiliser une base "catalogue" appelée Recovery CATalog (RCAT) dans laquelle il stocke les informations concernant les données sauvegardées. Les données sauvegardées peuvent être envoyées directement sur disque ou sur bande via un Media Manager c'est-à-dire un logiciel de sauvegarde tiers.

RPO : Recovery Point Objective : Perte de données maximale admissible sur incident.

RSC : Réseau Sans Couture / Seamless network.

RTO : Recovery Time Objective : Délai maximal d'interruption admissible après lequel les systèmes, applications, ou les activités doivent être rétablis après une interruption.

SAN : voir Storage Area Network.

SAN Media serveur : voir Media Serveur.

SAS : Serial Attached SCSI : Norme décrivant une nouvelle interface pour les disques durs, constituant une évolution des bus SCSI en termes de performances.

Scalabilité : Aptitude d'une application à maintenir son niveau de performance face à une augmentation de la charge, par l'augmentation de la capacité des ressources hardware. On distingue la scalabilité verticale (augmenter la capacité d'une ressource) et horizontale (augmenter le nombre de ressources).

SCSI : Small Computer System Interface : Standard définissant un bus permettant de connecter un périphérique à un ordinateur, et leur protocole de communication.

Service dégradé : Situation dans laquelle l'entreprise n'est pas en mesure de fournir ses prestations en mode nominal.

Server Less : Méthode de sauvegarde n'ayant aucun impact sur le serveur de production en utilisant un serveur tiers (c'est-à-dire autre que le serveur de sauvegarde ou l'application cliente) pour absorber la charge de mise sur bandes. Cette méthode est particulièrement utilisée pour des serveurs de base de données à forte volumétrie dont l'application ne permet aucun ralentissement du serveur et aucun arrêt de service. Elle impose d'avoir un serveur de manœuvre et de l'espace de stockage disponible pour effectuer la recopie intra baie.

SI : Système d'Information : ensemble organisé de ressources matérielles, logicielles et humaines, de données et de procédures qui permet de stocker, de traiter et de diffuser de l'information.

SLA : Le Service Level Agreement, est un document qui définit la qualité de service requise entre un prestataire et un client. Le Service Level Agreement, que l'on pourrait traduire en français par Contrat de niveau de service consiste donc en un contrat (ou la partie du contrat de service) dans lequel on formalise la qualité du service en question. Dans la pratique, le terme SLA est quelque fois utilisé en référence au temps de délivrance et/ou à la performance (du service) tel que définit dans le contrat.

SMS : Short Message Services : Message alphanumérique court (limité à 160 caractères) échangé entre téléphones portables et/ou fixes.

Snapshot : C'est une photographie, à un moment précis, de l'état d'une base de données, baie de disques ou d'une machine virtuelle VMware.

SRDF : Symmetrix Remote Data Facility, est un produit de la société EMC Corporation permettant la réplication de données informatiques entre 2 baies de disques en utilisant le mécanisme miroir Maître/Esclave. Les disques maîtres (nommé R1) sont en accès RW (Read-Write) et les disques esclaves (nommé R2) sont en WD (Write Disable). Plusieurs options de réplication existent : les plus connues étant le mode synchrone et le mode "adaptive copy" (asynchrone).

Storage Area Network (ou SAN) : réseau basé sur le protocole FC permettant de connecter des systèmes de stockage, des périphériques de sauvegarde et des serveurs via des cartes HBA.

Storage node : Serveur accédant directement à des lecteurs de bandes via le SAN et ayant des droits privilégiés de sauvegarde au niveau TiNa (ou NetWorker). Il s'agit le plus souvent d'un serveur applicatif à forte volumétrie ou nécessitant un débit de sauvegarde/restauration important.

STU : SStorage Unit : destination de stockage des sauvegardes NBU de type disque ou bande.

Time Navigator (ou TiNa) : Logiciel de sauvegarde de la société Atempo, concurrent du produit NetBackup.

VCB : VMware Consolidated Backup est la solution de proxy de sauvegarde fournie par VMware pour s'interfacer avec les outils de sauvegardes afin de permettre la sauvegarde et la restauration des Virtual Machines.

VLAN : Virtual Local Area Network : Un réseau local virtuel est un réseau informatique regroupant un ensemble de machines de façon logique (par port, par adresse MAC ou adresse IP) et non physique.

VTD : Un Virtual Tape Drive est l'émulation d'un lecteur de bandes dans une VTL.

VTL : Une Virtual Tape Library est un système de stockage informatique incluant un serveur, une grappe de disques et un logiciel capable d'émuler cet espace disque en bande magnétique. Le VTL vise à compenser les inconvénients du stockage sur bande par l'intermédiaire d'une couche (éventuellement intermédiaire) en système disque. Dans le cas du CentricStor, le disque ne sert que de tampon avant mise sur bande physique. Dans le cas du DXi, le stockage n'est constitué que par du disque (bien que l'externalisation sous forme de bande physique reste possible en ajoutant une robotique physique).

WWN : World Wide Name : identifiant unique d'un équipement FC ; équivalent au SAN de ce qu'est l'adresse MAC pour Ethernet.

Zone privée : Zone non accessible directement depuis les réseaux publics et inversement. Elle contient l'intelligence des services apportés par les projets (applications métier).

Zone trusted ou secure : Zone non accessible directement depuis les réseaux publics et inversement. Elle contient les données des projets (bases de données).

Zone publique : Zone en frontal des réseaux publics IP. Elle filtre les flux en provenance ou à destination de ces réseaux, et relaie si nécessaire les flux vers la zone privée. Elle peut être qualifiée de zone de présentation ou d'exposition du service. Elle ne contient aucune donnée.

Zoning : Autorisation de communication d'au moins deux périphériques sur le réseau SAN. Comparable à un VLAN dans le monde IP.

ZSV : Zone de SauVegarde : la DMZ dans un IAS qui héberge l'infrastructure de sauvegarde de cet IAS. Cf. annexe I

Table des tableaux

Tableau I : matrice des activités de la MOE sauvegarde	12
Tableau II : Les projets du programme MIO.....	16
Tableau III : répartition des agents TiNa installés	43
Tableau IV : infrastructure actuelle.....	43
Tableau V : environnement technique	44
Tableau VI : packages Platon à déployer	44
Tableau VII : contraintes exprimées dans le PMP	49
Tableau VIII : répartition des clients par media serveur	55
Tableau IX : répartition temporelle des sauvegardes	58
Tableau X : comparatif des robotiques	60
Tableau XI : taille des 19 catalogues TiNa	66
Tableau XII : estimation de la taille des 2 catalogues NBU	66
Tableau XIII : justification du passage en CVAT	77
Tableau XIV : comparaison du mode de sauvegarde des bases Oracle	81
Tableau XV : macro-calendrier / jalons exprimés dans le PMP	89
Tableau XVI : espace et consommation électrique des éléments de l'infrastructure NBU	91
Tableau XVII : implantation physique de l'infrastructure NBU	93
Tableau XVIII : adressage IP de l'infrastructure NBU	94
Tableau XIX : suivi du déploiement de l'infrastructure NBU	95
Tableau XX : synthèse globale.....	104
Tableau XXI : gestion des risques	104
Tableau XXII : bilan pilotage projet	105
Tableau XXIII : gains.....	105
Tableau XXIV : décommissionnement des serveurs Tina.....	106
Tableau XXV : taux de réussite des jobs de sauvegarde	108

Table des illustrations

Illustration 1 : la marque Orange dans le monde	7
Illustration 2 : mon positionnement dans l'organigramme FT.....	9
Illustration 3 : accès au mail Orange depuis son téléphone portable : via la plate-forme MDSP !....	14
Illustration 4 : technologies présentes sur la plate-forme MDSP.....	15
Illustration 5 : interactions entre les projets du programme MIO	17
Illustration 6 : sauvegarde LAN	21
Illustration 7 : sauvegarde SAN ou LAN Free	22
Illustration 8 : sauvegarde NDMP.....	23
Illustration 9 : contraintes pour garantir l'intégrité des données.....	25
Illustration 10 : RPO et RTO	26
Illustration 11 : gains sur les processus de reprise	27
Illustration 12 : principe d'implantation du CentricStor.....	29
Illustration 13 : principe d'implantation du DXi	29
Illustration 14 : principe de la déduplication.....	30
Illustration 15 : exemple de déduplication à la cible.....	32
Illustration 16 : les six facteurs clés de succès.....	35
Illustration 17 : les jalons TTM.....	36
Illustration 18 : organigramme de la division ROSI	38
Illustration 19 : organigramme de la direction DPS.....	39
Illustration 20: organigramme de la direction DOE.....	40
Illustration 21 : schéma logique détaillant les flux de production et d'administration.....	46
Illustration 22 : schéma prévisionnel d'intégration des applications MDSP dans les IAS	47
Illustration 23 : fabric étendue du SAN de backup en production	48
Illustration 24 : schéma fonctionnel d'une sauvegarde TiNa.....	52
Illustration 25 : schéma fonctionnel d'une sauvegarde NBU	53
Illustration 26 : scénarios robotiques	61
Illustration 27 : évolution des fabrics de l'IAS	63
Illustration 28 : intégration dans le réseau IAS	68
Illustration 29 : scénario 1 d'architecture SAN.....	70
Illustration 30 : scénario 2 d'architecture SAN.....	72
Illustration 31 : scénario 3 d'architecture SAN.....	73
Illustration 32 : sur Vélizy, double attachement réseau d'administration du master et triple attachement du DXi	74
Illustration 33 : schéma d'architecture.....	76
Illustration 34 : solution actuelle : RMAN Server Less + TiNa + STK.....	79
Illustration 35 : ordonnancement des sauvegardes Server Less.....	80
Illustration 36 : RMAN Server Less et RMAN Lan Free	81
Illustration 37 : solution cible : RMAN LAN Free + NBU + DXi	82
Illustration 38 : phase 0 : état des lieux.....	85
Illustration 39 : phase 1 : déploiement.....	86
Illustration 40 : phase 2 : migration	87
Illustration 41 : phase 3 : fin du projet.....	88
Illustration 42 : roadmap	89
Illustration 43 : architecture mise en place	103
Illustration 44 : définition de la chaîne de soutien	107
Illustration 45 : qualité de service	109
Illustration 46 : suivi des DXis	109
Illustration 47 : projection des consommations d'un DXi	110