

Chapitre 6

Vision Informatique Technique Architecture Physique

NFE107



Chapitre 6

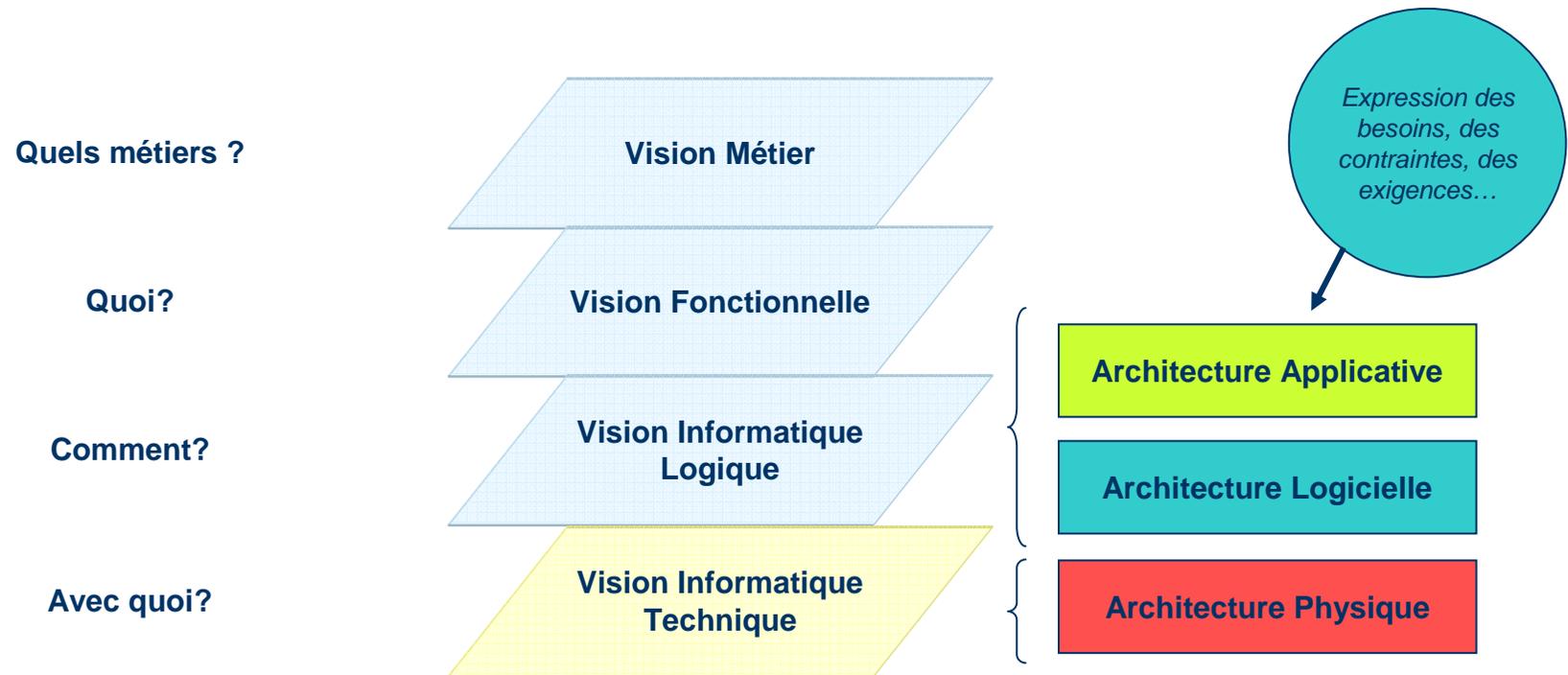
Vision Informatique Technique Architecture Physique

6.1 Introduction



Positionnement de la Vision Informatique

- Urbanisation / Architecture



Introduction

- Objectifs :
- Architecture Technique
 - elle structure les moyens d'infrastructure technique à mettre en œuvre pour informatiser l'activité de l'entreprise
 - elle décrit et organise
 - les différents moyens matériels (serveur, poste client ...)
 - les logiciels de base (systèmes d'exploitation, SGBD, AGL ...)
 - les moyens de communication entre elles (réseaux, middleware...)

Elle répond à la question AVEC QUOI ?

Une démarche en 6 étapes

- De manière itérative et incrémentale l'architecte va dans la phase d'architecture physique :
 1. Décrire l'architecture physique pour chaque bloc applicatif, en précisant les moyens d'infrastructure technique à mettre en œuvre (implantation des composants, topologie) :
 - Plates-formes matérielles (poste de travail, serveurs départementaux, serveurs d'entreprises, stockage, ...)
 - Logiciels de base (système d'exploitation, SGBD, middleware, serveurs d'applications, annuaires...)
 - Réseaux et télécommunications (réseau locaux (LAN) ou longues distances (WAN), débit, switch, routeurs, proxies, firewalls, ...)
 2. Définir la qualité de service attendue pour le système à partir des contraintes exprimées et décrire les solutions à mettre en œuvre :
 - Performance
 - Disponibilité
 - Sécurité

Une démarche en 6 étapes (suite)

3. Définir un modèle de dimensionnement théorique à partir des hypothèses et des contraintes existantes et exprimées, permettant de dimensionner :
 - les moyens matériels
 - les logiciels (OS, SGBDR, ...)
 - les réseaux (calcul de bande passante)
4. Piloter/participer à l'étude de prototypage et au benchmark permettant ainsi de valider/affiner :
 - le modèle de dimensionnement de l'infrastructure technique
5. Définir l'architecture de l'exploitation (administration / supervision du SI)
 - outils de supervision
 - procédures de déploiement et de configuration, procédures d'exploitation, de supervision
6. Déterminer le coût du système

Coûts

- Points à prendre en compte lors de l'estimation des coûts
 - Ressources humaines
 - Formation
 - Développement
 - Administration (exploitation / supervision)
 - Logiciels / Progiciels (licences, coûts récurrents, support) : catalogues et tarifs éditeurs
 - Matériels (serveurs, réseau) : catalogues et tarifs constructeurs
 - Télécommunications : catalogues et tarifs opérateurs
 - Assistance : plateforme d'assistance / support (help desk)

Chapitre 6

Vision Informatique Technique Architecture Physique

6.2 Architecture Physique



Architecture Physique

- La description de l'architecture physique, de chacun des blocs applicatifs, présentant les moyens d'infrastructure technique à mettre en oeuvre (implantation des composants, topologie) est réalisée à partir :
 - De la cartographie applicative des flux,
 - Du découpage en couches (layer view) logicielles et
 - Du découpage vue en niveaux (tier view)
- Fournir un schéma présentant pour chaque bloc applicatifs l'architecture physique générale faisant apparaître tous les composants de l'architecture physique de production
 - Infrastructure logicielle :
 - logiciels de base (système d'exploitation, SGBD, middleware, serveur web, serveurs d'applications, annuaires, SSO, ...)
 - Infrastructure matérielle :
 - plates-formes matérielles (poste de travail, serveurs départementaux, serveurs d'entreprises, stockage...)
 - réseaux et télécommunications : réseau locaux (LAN) ou longues distances (WAN), débit, plan d'adressage IP, switch, routeurs, proxies, firewalls, ...
- Décrire de façon détaillée, pour chacun des blocs applicatif, les configurations de :
 - L'infrastructure logicielle
 - systèmes d'exploitation, socle logiciel, progiciel, configuration, ...
 - L'infrastructure matérielle
 - modèles, configuration RAM, disque, CPU, réseau, ...
- Lors de la première itération, le nombre de composants (matériels, logiciels, réseaux) constituant l'infrastructure logicielle et matérielle n'est pas encore connu, l'architecte doit procéder de façon itérative et incrémentale

Schéma d'Architecture Physique Infrastructure logicielle

- Exemple d'infrastructure logicielle J2EE

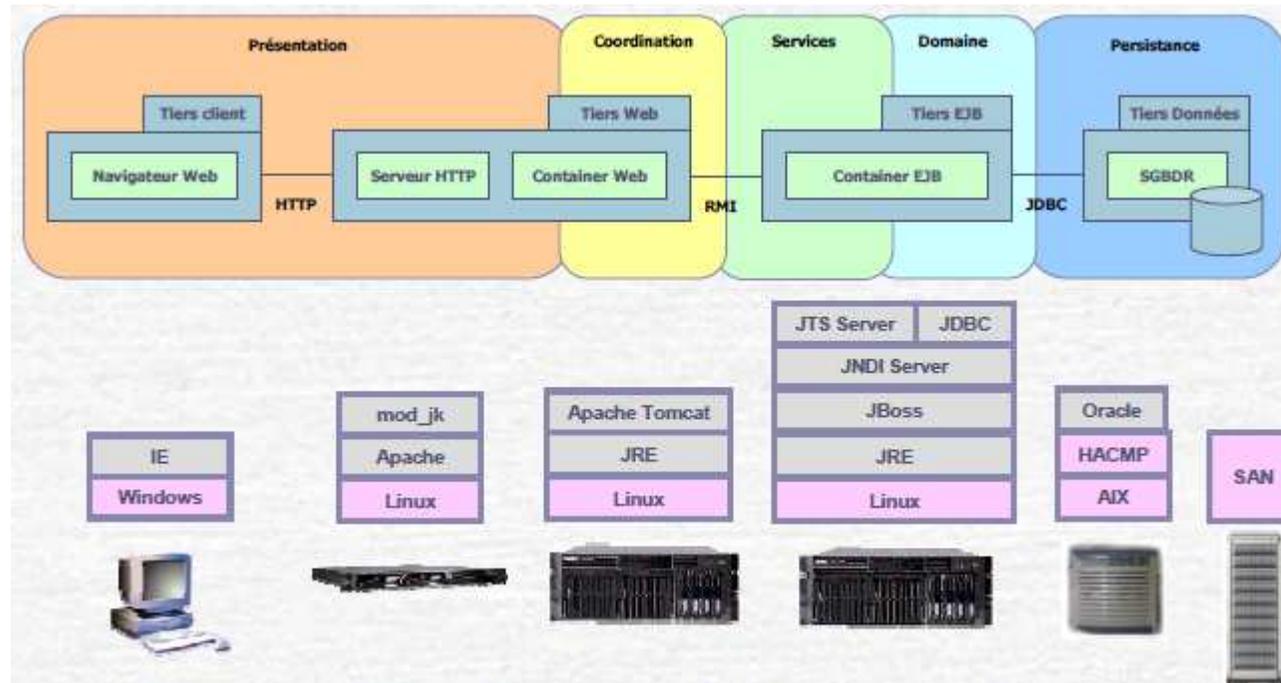
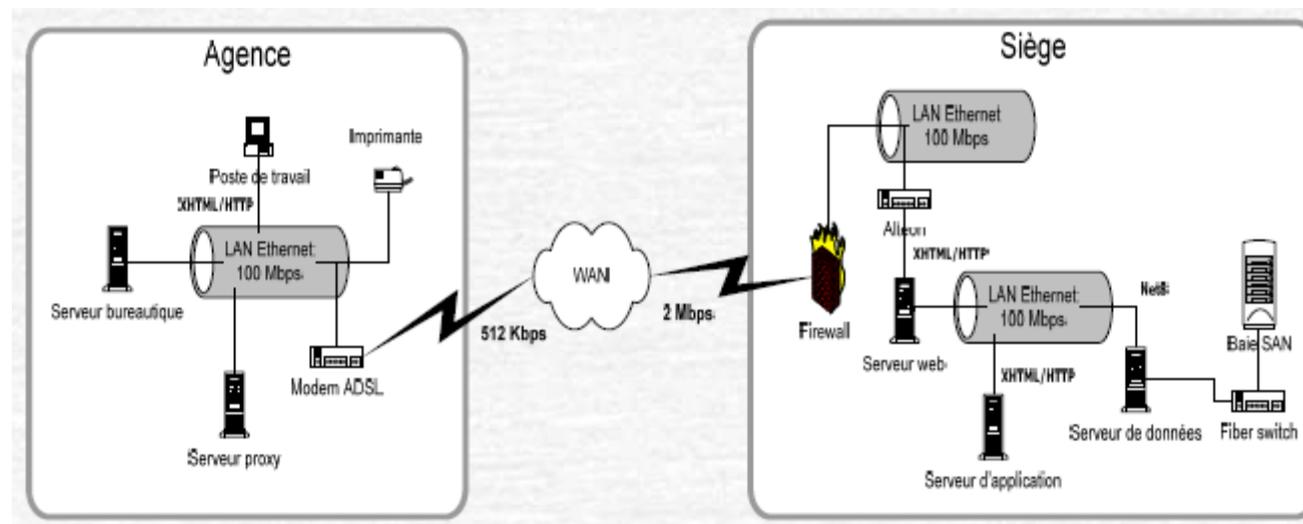


Schéma d'Architecture Physique Infrastructure matérielle

- Exemple d'infrastructure matérielle



Chapitre 6

Vision Informatique Technique Architecture Physique

6.3 Qualité de service



Contrat de qualité de service

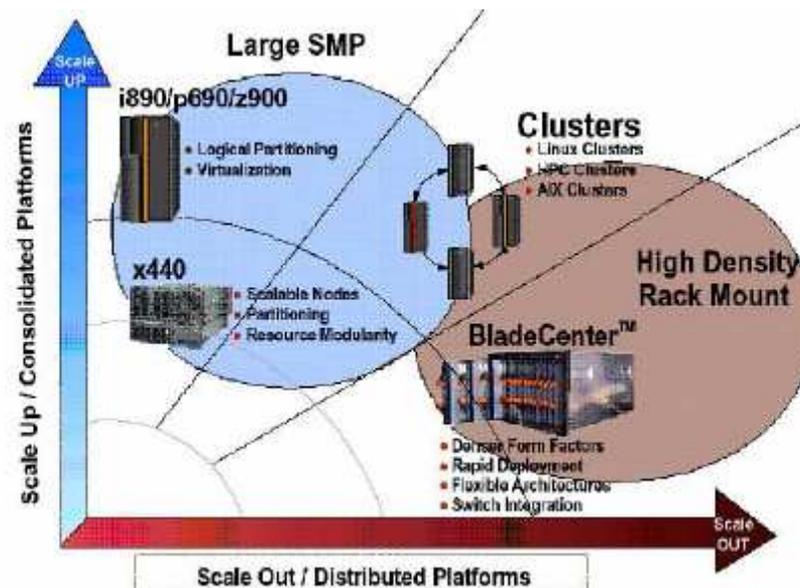
- Le contrat de qualité de service (Quality of Service) définit certains critères permettant d'apprécier :
 - L'utilisation du système
 - L'efficacité demandée au système
- Les critères sont les suivants :
 - Performance
 - Disponibilité
 - Sécurité
- Des contrats de qualité de service doivent être mis en place entre les différents blocs applicatifs afin de garantir le niveau de service attendu

Performances

- Le critère concernant les performances attendues par le système est un élément très structurant du modèle de dimensionnement
- Les aspects les plus importants à prendre en compte sont les suivants :
 - TP : débit transactionnel, temps de réponse
 - Batch : durée des traitements batchs
 - Montée en charge (scalabilité)
 - Mécanismes de cache (navigateur, proxy, serveur web, cache serveurs d'applications, cache base de données...)
 - Gestion de l'affinité de session
 - Réplication de données, dénormalisation du modèle de données...
- L'architecte doit préconiser des solutions permettant de maximiser les performances du système
- Les performances sont très dépendantes du dimensionnement matériel et réseau

Montée en charge (scalabilité)

- La « scalabilité » se définit comme la capacité du système à :
 - faire face aux montées en charge
 - évoluer par « simple » ajout de matériel
- « Scalabilité » horizontale et verticale (exemple IBM)



Disponibilité

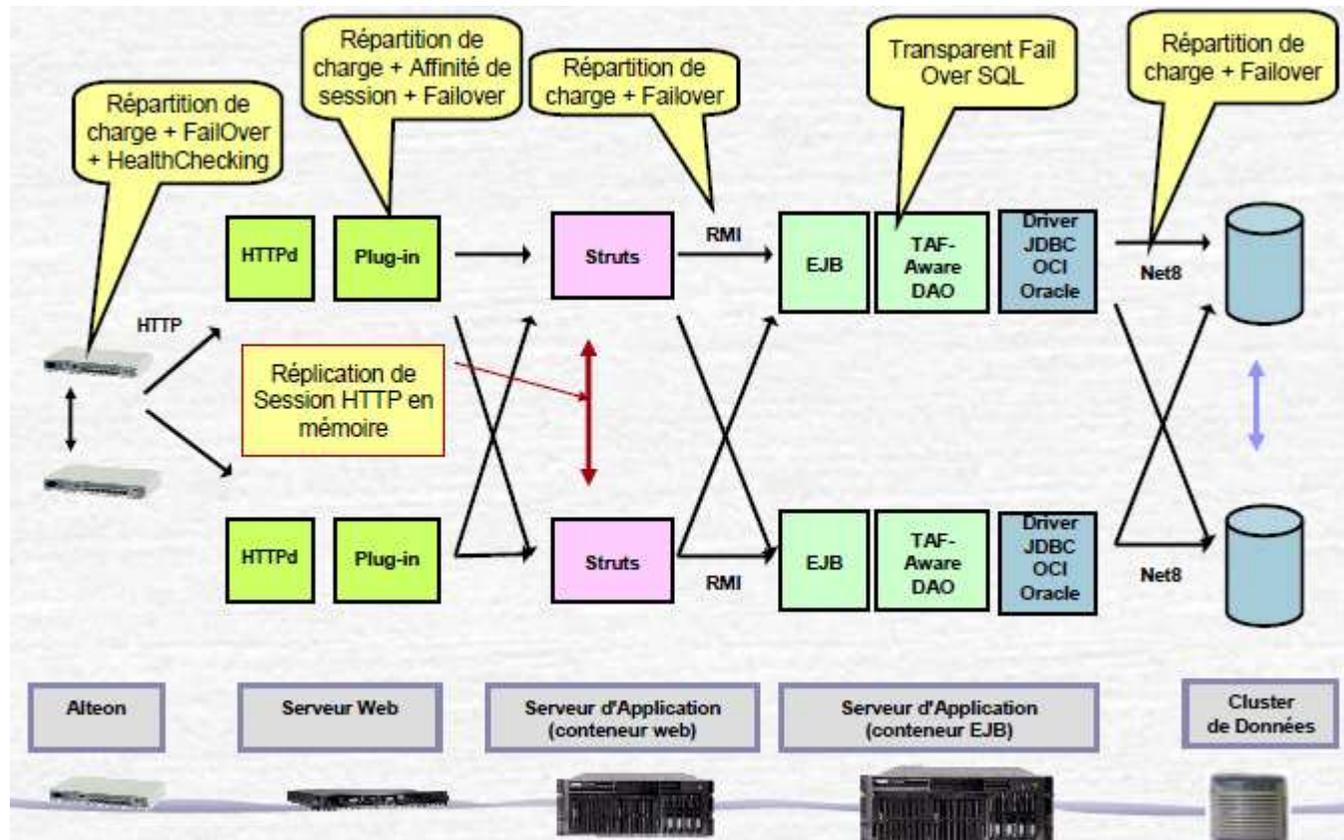
- Le critère de disponibilité et de continuité de service du système adresse les problématiques suivantes :
 - niveau de disponibilité du service :
 - durée maximale des temps d'indisponibilité
 - fréquence maximale des indisponibilités
 - plages d'indisponibilité éventuelles (exemple : travaux d'exploitation programmés ou exceptionnels):
 - sauvegardes, maintenance serveur, maintenance réseau, ...
- La réponse à ces problématiques passe par la mise en œuvre d'un certain nombre de mécanismes et de solutions :
 - haute disponibilité (répartition de charge et fail-over)

Avail-ability (%)	Cumulative Outage per Year
99.9999	32 sec
99.999	5 min
99.99	53 min
99.9	8.8 hrs
99	87 hrs. (3.6 days)
90	876 hrs. (36 days)

Haute disponibilité

- La haute disponibilité comporte deux axes :
 - La répartition de la charge
 - La gestion du fail-over (redondance matérielle des machines, mise en place de cluster (grappes actif/passif, actif/actif) de serveurs et base de données, utilisation pour les espaces de stockage de RAID, SAN, NAS)
- Répartition de charge (load-balancing)
 - Possible à plusieurs niveaux dans l'architecture physique
 - Plusieurs solutions existent sur le marché :
 - Matériel : boîtier (exemple : alteon), serveur DNS (round-robin), ...
 - Logiciel : serveur web, serveur d'application, ...
 - La mise en œuvre d'une architecture SOA utilisant des services sans état (stateless) permet de faciliter la répartition de charge
- Fail-over
 - Mise en œuvre de cluster (grappes actif/passif, actif/actif)
 - Réplication du contexte de session
 - Mémoire ou persistant
- La mise en œuvre de mécanismes de haute disponibilité nécessite la mise en œuvre synchronisation des horloges des serveurs via NTP avec une base de temps fournie par une horloge atomique

Exemple de mise en œuvre de la haute disponibilité avec J2EE



Répartition / Synchronisation

- Dans le but d'optimiser les performances, il est parfois nécessaire de dupliquer de l'information
 - Stratégie de réplication / synchronisation des données
 - Référentiel maître / esclave
 - Architecture distribuée / répartie
 - Problématique d'habilitation et de juridiction...
- Afin de garantir la continuité de service (24/24, 7/7), il peut être envisagé de laisser une instance de la base de données en consultation seule, pendant qu'une autre instance exécute des traitements batchs. Par la suite une bascule et une synchronisation des données seront réalisées

RAID

- RAID (Redundant Array of Independent Disks)
 - Au lieu de stocker les informations sur un seul gros disque, on les mémorise sur plusieurs petits disques.
 - Cela permet en particulier de répartir à plusieurs endroits des informations cruciales (des index par exemple), d'accélérer les accès aux données en pratiquant l'entrelacement des informations (des parties consécutives d'un fichier sont placées sur des disques différents ce qui permet de les récupérer toutes en même temps au lieu de les lire les unes à la suite des autres).
 - La bande passante des entrées sorties sur disque est ainsi considérablement améliorée. Enfin, après un crash d'un des disques, la récupération des données est facilitée. Plusieurs niveaux de disques Raid sont disponibles sur le marché : entrelacement des informations sur les disques (RAID niveau 0), duplication en miroir des fichiers (RAID niveau 1), ...

SAN

- SAN (Storage Area Network)
 - On réunit sous cette dénomination un réseau de périphériques de stockage, reliés les uns aux autres par des canaux de type SCSI (Small Computer System Interface), SSA (Serial Storage Architecture), ESCON (Enterprise System Connection), ou Fibre Channel.
 - On peut réaliser un réseau SAN de manière centralisée ou décentralisée.
 - SAN centralisé : les différentes unités centrales ont accès à une seule unité de stockage, constituée généralement d'un ensemble de disques Raid (Redundant Array of Independent Disks).
 - SAN décentralisé : les différentes unités de traitement ont accès à plusieurs unités de stockage.

NAS

- NAS (Network Attached Storage)
 - Cette appellation désigne une unité de stockage reliée au reste du système par des protocoles de communication classiques (Ethernet, TCP/IP, etc.), exactement comme tout autre serveur de fichiers.
 - Le dispositif NAS ne dispose pas de son propre système d'exploitation mais utilise un micro noyau dédié à la gestion des entrées-sorties comme NFS (environnement Unix), SMB/CIFS (DOS/Windows) ou NCP (Novell Netware).

Prévention des sinistres

Reprise sur incident

- La reprise sur incident nécessite d'adresser les points suivants dans le cadre de l'architecture physique afin de permettre de redémarrer l'activité de l'entreprise dans les meilleurs délais et avec le moins de perte d'information possible :
 - définir une stratégie de réplication des différents sites
 - définir un ou des sites de secours
 - définir un PRA (Plan de Reprise d'Activité)
 - processus de basculement sur le site de secours
 - processus de retour à la normal
 - réaliser des tests du plan de secours (PRA) à fréquence régulière

Sécurité

- Identification et Authentification
 - Décrire les mécanismes proposés pour l'identification et l'authentification
 - Authentification basique (identifiant / mot de passe), annuaire, certificats X.509, SSO
- Contrôles d'accès
 - Décrire les mécanismes proposés pour contrôler les accès
 - Firewall, DMZ, reverse proxy, ...
- Habilitations et Profil
 - Décrire les mécanismes proposés pour gérer les profils et les habilitations
 - Profil utilisateur
 - Habilitations applicatives
- Confidentialité
 - Décrire les mécanismes permettant de garantir la confidentialité des données
 - Chiffrement de données
- Intégrité
 - Décrire les mécanismes proposés pour garantir l'intégrité des données
 - Calcul de hash (MD5, ...)
- Non répudiation
 - Décrire les mécanismes permettant d'assurer la non répudiation
 - Signature des données

Chapitre 6

Vision Informatique Technique Architecture Physique

6.4 Démarche de dimensionnement



Démarche de dimensionnement

1. Créer un modèle de dimensionnement théorique à partir des hypothèses de :
 - Dimensionnement statique
 - Dimensionnement dynamique
2. Utiliser des métriques et des abaques éditeurs et constructeurs pour faire un premier dimensionnement du système à partir des éléments du modèle de dimensionnement théorique
3. Développer un prototype implémentant des cas représentatifs de l'utilisation du système
4. Mettre en œuvre un benchmark (sélection de scénario représentatifs de l'utilisation du système, constitution de jeux d'essai, ...)
 - le benchmark permettra de valider et d'affiner le modèle de dimensionnement théorique
5. Comparer les résultats du benchmark avec le modèle de dimensionnement théorique
6. Affiner avec l'ensemble des éléments le dimensionnement de la plateforme de production
7. Préciser sur le schéma de l'architecture physique détaillée le nombre de machines et les débits réseaux
8. Ne pas oublier d'identifier les possibilités de mutualisation des infrastructures (serveurs, stockage, ...)

Prototype et benchmark

- La mise en œuvre d'un prototype opérationnel implémentant un ou plusieurs cas d'utilisation (use-case) significatifs, permet de :
 - valider les choix d'architecture définis dans le dossier d'Architecture (validation de la faisabilité technique)
 - éprouver le socle technique (outils, frameworks, ...) sur une réalisation concrète
 - valider la méthodologie sur l'ensemble de la chaîne de production (des spécifications techniques aux procédures de déploiement)
- Un plan de performances appelé benchmark est réalisé à partir du prototype. Il permettra de valider et d'affiner le modèle de dimensionnement théorique
- Etapes du benchmark
 1. Sélection de scénario représentatifs de l'utilisation du système (définition de profils)
 2. Constitution de jeux d'essai (échelle 1 ou utilisation éventuelles d'homothéties)
 3. Sélection de l'outil de charge (injecteur, contrôleur), mise en place de sondes logicielles et matérielles
 4. Tests (TP et batchs) en charge et en endurance
 5. Analyse des résultats
 6. Phase d'optimisation

Éléments de volumétries

- Volumétrie Statique
 - Nombre d'enregistrements
 - Taille d'enregistrement (octets)
 - Historisation (rétention, archivage)
 - Estimation du volume total des données fonctionnelles, y compris les données historisées.
 - Volume physique total...
- Volumétrie Dynamique
 - Nombre d'utilisateurs du système (internes, externes et partenaires)
 - Nombre de transactions simultanées
 - Batches
 - Nombre
 - Fréquence
 - Durée
 - Estimation du nombre de connexions SGBD simultanées
 - Estimation du nombre de requêtes SQL (profils de requêtage)
 - Ratios écriture/lecture...

Modèle de dimensionnement théorique

- A partir de l'expression des besoins le système doit être dimensionné afin d'atteindre les performances attendues
- Le modèle de dimensionnement théorique se base sur les hypothèses de dimensionnement statique et dynamique, il permet de déterminer les éléments suivants :
 - Machines (poste de travail, serveurs)
 - Nombre de transaction par seconde
 - Durée et volumétrie des traitements batchs
 - Réseau local (LAN) et distant (WAN)
 - Estimation bande passante réseau, temps de latence
 - Profils de consommation de bande passante
 - Espace de stockage (disque, NAS, SAN)
 - Volumétrie des données
 - Espace disque
 - Nombre de disque (niveau de RAID)
- A partir de ces éléments, des métriques et des abaques éditeurs et constructeurs sont utilisés pour réaliser un premier dimensionnement du système

Chapitre 6

Vision Informatique Technique Architecture Physique

6.5 Exploitation (Administration &
Supervision)



Les environnements du SI

- L'architecte va définir et dimensionner les plateformes :
 - de développement
 - d'intégration
 - de benchmark
 - de recette
 - de formation
 - de pré-production / qualification
 - de production

Exploitation

- L'exploitation prend en compte les points suivants :
 - Organisation
 - Acteurs
 - Planification
 - Sécurité des données
 - Stratégie globale
 - Sauvegardes
 - Restaurations
 - Archivages, purge
 - Bascule site de secours
 - Administration
 - Supervision
 - Editions / impressions
 - Assistance utilisateurs

Administration

- Prise de contrôle à distance
- Télédistribution
- Ordonnancement des traitements (notamment des batches)
- Sauvegarde / Restauration
 - L'architecture de sauvegarde / restauration du système est un élément très structurant à prendre en compte.
 - Les points suivants sont à adresser :
 - Dispositif de sauvegarde: serveur de sauvegarde, réseau de sauvegarde, CD, DVD, DAT, DLT, SDLT, LTO (Linear Tape Open)...
 - Outils (logiciel) : IBM Tivoli Storage Manager, Atempo TiNa (Time Navigator), ...
 - Procédure d'archivage
 - Procédure de sauvegarde / restauration
 - Stratégie de sauvegarde / restauration
- Afin de ne pas avoir de mauvaise surprise, il est primordial de faire des tests de restauration à intervalle régulier

Supervision

- La supervision du SI permet d'assurer aux utilisateurs le niveau de qualité de service convenu
- La supervision comprend :
 - le pilotage de l'exploitation comprenant la surveillance de :
 - l'infrastructure matérielle : serveurs et réseaux
 - l'infrastructure logiciel de base : OS, middleware, SGBD, annuaire, serveurs d'applications
 - Infrastructure applicative : développement spécifiques permettant de remonter les anomalies au système de supervision (logs, syslog, trap SNMP, ...)
 - la supervision technique :
 - Suivi de la volumétrie des données
 - Gestion de la trace
 - Statistiques et logs
 - Mesure de la qualité de service fournie par rapport au niveau de service attendu (robot d'appel technique et fonctionnel)
 - Déclenchement des plans de secours en cas de dysfonctionnements graves
 - Capacity planning : prévision d'évolution des infrastructures de production en fonction de l'évolution de leur utilisation (augmentation du nombre d'utilisateurs, augmentation de la taille des fichiers gérés ...)