

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
UNIVERSITE MOULOUD MAMMERI DE TIZI-OUZOU



FACULTE DU GENIE ELECTRIQUE ET D'INFORMATIQUE
DEPARTEMENT D'INFORMATIQUE

Mémoire de Fin d'Etudes
De MASTER PROFESSIONNEL
Domaine : Mathématiques et Informatique
Filière : Informatique

Spécialité : Ingénierie des Systèmes d'Information

Présenté par
BOUDJEDAIMI Yasmine
YEFSAH Lila

Thème

**Mise en place d'une plateforme de
supervision de bases de données**

Mémoire soutenu publiquement le 16/09/2018 devant le jury composé de :

Présidente : Mme FELLAG SAMIA
Encadreur : Mr SI MOHAMMED MALIK
Examinatrice : Mme SINI GHENIMA

Remerciements

Nous exprimons notre profonde gratitude et respectueuses reconnaissances à notre encadreur **Mr Si Mohammed Malik** pour sa bonne volonté d'accepter de nous encadrer, pour tout le temps qu'il nous a octroyé et pour tous les conseils qu'il nous a prodigués.

Nous remercions nos encadreurs de Société Général Algérie pour leur suivi et leur sympathie.

Que les membres du jury trouvent ici nos remerciements pour avoir accepté d'examiner notre travail.

Des remerciements également à toute l'équipe pédagogique ainsi que les intervenants professionnels extérieurs et les enseignants de l'Université Mouloud MAMMERI pour toutes les connaissances théoriques qu'ils nous ont apportées tout au long de notre cursus universitaire.

Remerciements particuliers à nos familles respectives pour leur soutien constant et leurs encouragements.

Table des matières

INTRODUCTION GENERALE	9
Problématique	10
Chapitre I : Généralités	11
1. Introduction	11
2. Organisme d'accueil.....	11
2.1. Présentation de l'organisme d'accueil	11
2.2. Activités de SGA.....	13
2.2.1 Activité auprès des entreprises.....	13
2.2.2 Global Transaction Banking (GTB).....	13
2.2.3 Le Marché des Particuliers	13
2.2.4 Le Marché des Professionnels	14
2.3. Présentation de la DSI.....	15
2.3.1 Organigramme.....	15
2.3.2 Missions.....	16
3. Généralités sur les bases de données	17
3.1. Définition.....	17
3.2. Objectifs.....	17
3.3. Les systèmes de gestion de bases de données (SGBD)	18
3.3.1 Définition	18
3.3.2 Objectifs.....	18
3.3.3 Fonctionnalités	19
3.3.4 Conception d'une base de données	19
3.3.5 Les systèmes de gestion de bases de données relationnelles	20

3.4. Classification des bases de données	21
3.4.1 Classification selon l'usage.....	21
3.4.2 Classification selon l'emplacement.....	21
4. Base de données Informix	22
4.1. IBM Informix Database Server.....	22
4.2. IBM Informix database.....	22
4.3. Architecture du serveur de base de données Informix.....	22
4.4. Informix vs Oracle.....	24
4.5. Caractéristiques.....	26
4.6. Informix et le Cloud.....	26
5. Conclusion	27
Chapitre II : La supervision de bases de données	28
1. Introduction	28
2. La supervision	28
2.1. Définition.....	28
2.2. Principe.....	28
2.3. Types.....	29
2.4. Aspects.....	30
2.5. Quelques Standards normalisés.....	30
2.6. Intérêt	33
3. Supervision de bases de données	33
3.1. Définition.....	33
3.2. Métriques de supervision.....	33
3.3. Surveillance des performances de bases de données	34
3.3.1 Objectifs.....	35
3.3.2 Avantages.....	35

3.4. Type d'outils de surveillance des bases de données.....	35
3.5. Quelques outils de supervision.....	36
3.6. Les contraintes de la supervision de bases de données.....	42
4. Conclusion.....	42
Chapitre III : Conception.....	43
1. Introduction.....	43
2. Contexte du projet.....	43
3. Schéma de la base de données supervisée.....	44
4. Architecture de la solution.....	45
5. Présentation des technologies utilisées.....	45
5.1. InfluxDB.....	45
5.1.1 Définition.....	45
5.1.2 Principales caractéristiques.....	46
5.1.3 Concepts clés.....	46
5.2. Grafana.....	47
5.2.1 Définition.....	47
5.2.2 Architecture.....	48
6. Démarche suivie.....	51
7. Conclusion.....	53
Chapitre IV : Réalisation.....	54
1. Introduction.....	54
2. Environnement de travail.....	54
2.1. Environnement matériel.....	54
2.2. Environnement logiciel.....	54
2.2.1 Système d'exploitation.....	54
3. Architecture de la plateforme.....	55

4. Fonctionnement	55
5. Installation et configuration.....	56
5.1. Installation de InfluxDB	56
5.2. Installation de Grafana.....	58
5.3. Mise à jour de Perl.....	60
5.4. Installation de DBD::Informix	60
6. Mise en place de la plateforme de supervision	63
6.1. Création de la base de données Influxdb	63
6.2. La collecte de métriques.....	63
6.3. Connexion Grafana influxdb	64
6.4. Création de visualisations.....	66
6.5. Mise en place d’alertes	70
7. Conclusion	72
Annexe.....	73
1. Installation de la base de données.....	73
1.1 Préparation de l'environnement	73
1.2 Installation du serveur de base de données informix IDS	73
1.3 Configuration post installation.....	79
Conclusion générale.....	81
Bibliographie	82

Liste des Figues

Figure 1: Organigramme de Société Générale Algérie.....	12
Figure 2: Organigramme de la Direction des Systèmes d'Information	15
Figure 3: Base de données	17
Figure 4: Architecture du serveur de base de données Informix	23
Figure 5: Fonctionnalités de Nagios.....	38
Figure 6: Centreon	39
Figure 7: Schéma de base de données.....	44
Figure 8: Architecture de la solution de supervision	45
Figure 9: Architecture de Grafana	48
Figure 10: Sources de données dans Grafana.....	50
Figure 11: Exemple de panel	51
Figure 12: Architecture de la solution	55
Figure 13: Connexion à Grafana	60
Figure 14 : Ajout d'une source de données.....	64
Figure 15: Configuration de la source de données.....	65
Figure 16: Test de la connexion à la source de données.....	65
Figure 17: Exemple de dashboard	66
Figure 18: Database size	67
Figure 19: Buffer,page and disk writes	68
Figure 20: Buffer, lock, checkpoint and latch waits	69
Figure 21: Physical and logical logging	70
Figure 22 : Configuration du canal de notification	71
Figure 23: Configuration du SMTP	71
Figure 24: Exemple d'alerte.....	72

Liste des tableaux

Tableau 1: Comparaison entre Informix et Oracle.....	25
--	----

INTRODUCTION GENERALE

L'informatique étant devenue l'épine dorsale de l'entreprise quel que soit son secteur d'activité, le système d'information est au centre de l'activité de différentes entités métiers et doit fonctionner pleinement et en permanence pour garantir l'efficacité de l'entreprise. A tous les niveaux, les réseaux, les terminaux utilisateurs, les serveurs d'applications et les données constituent autant de maillons sensibles dont la disponibilité et la qualité de service conditionnent le bon fonctionnement de l'entreprise. Les problèmes liés à l'informatique doivent donc être réduits au minimum, car une indisponibilité du système d'information a des impacts très préjudiciables sur l'activité et sur la notoriété d'une entreprise. [1]

C'est dans ce contexte que se situe l'objectif de notre projet qui se caractérise par la mise en place d'une plateforme de supervision des bases de données qui permettra à Société Générale Algérie d'avoir une vue globale de l'état de santé de ses serveurs de bases de données.

Le présent document est structuré en quatre chapitres. Dans le premier chapitre nous décrivons l'organisme d'accueil et présentons des généralités sur les bases de données. Ensuite, dans le deuxième chapitre nous abordons la supervision informatique en général et plus précisément la supervision des bases de données qui est l'objet de notre travail et nous présentons quelques outils de supervision existants. Le troisième chapitre est consacré à la conception nous présentons la démarche suivie pour la mise en place de notre plateforme de supervision et expliquons chacune des différentes étapes. Le quatrième chapitre sera réservé à la réalisation de notre application.

Problématique

Certaines parties du système d'information sont vitales pour l'entreprise, comme c'est le cas des serveurs de bases de données, l'administrateur de celles-ci doit maintenir en production la base de données en temps réel, garantir sa disponibilité et réduire l'impact des incidents sur celle-ci.

Cependant tous les incidents ne sont pas logés à la même enseigne en termes de criticité. L'administrateur doit surveiller les bases en permanence, une tâche qui peut s'avérer compliquée sans outil de surveillance adéquat vu le nombre important d'informations à analyser. C'est pour cela qu'il est impératif de se munir d'une solution de supervision pour faciliter le suivi de l'état de la base de données et ses performances.

Chapitre I : Généralités

1. Introduction

Dans ce premier chapitre nous commençons par faire une présentation de l'organisme d'accueil où nous avons effectué notre stage pratique qui est Société Générale Algérie puis nous présentons quelques généralités sur les bases de données.

2. Organisme d'accueil

2.1. Présentation de l'organisme d'accueil

Société Générale Algérie, détenue à 100% par le Groupe Société Générale, est l'une des toutes premières banques privées à s'installer en Algérie, soit depuis 2000.

Son réseau, en constante croissance, compte 87 agences dont 11 Centres d'Affaires ou business center dédiés à la clientèle des entreprises. Société Générale Algérie offre une gamme diversifiée et innovante de services bancaires à près de 400 000 clients particuliers, professionnels et entreprises. L'effectif de la banque est de 1 396 collaborateurs au 31 décembre 2016. [2]

La figure ci-dessous représente l'organigramme de Société Générale Algérie :

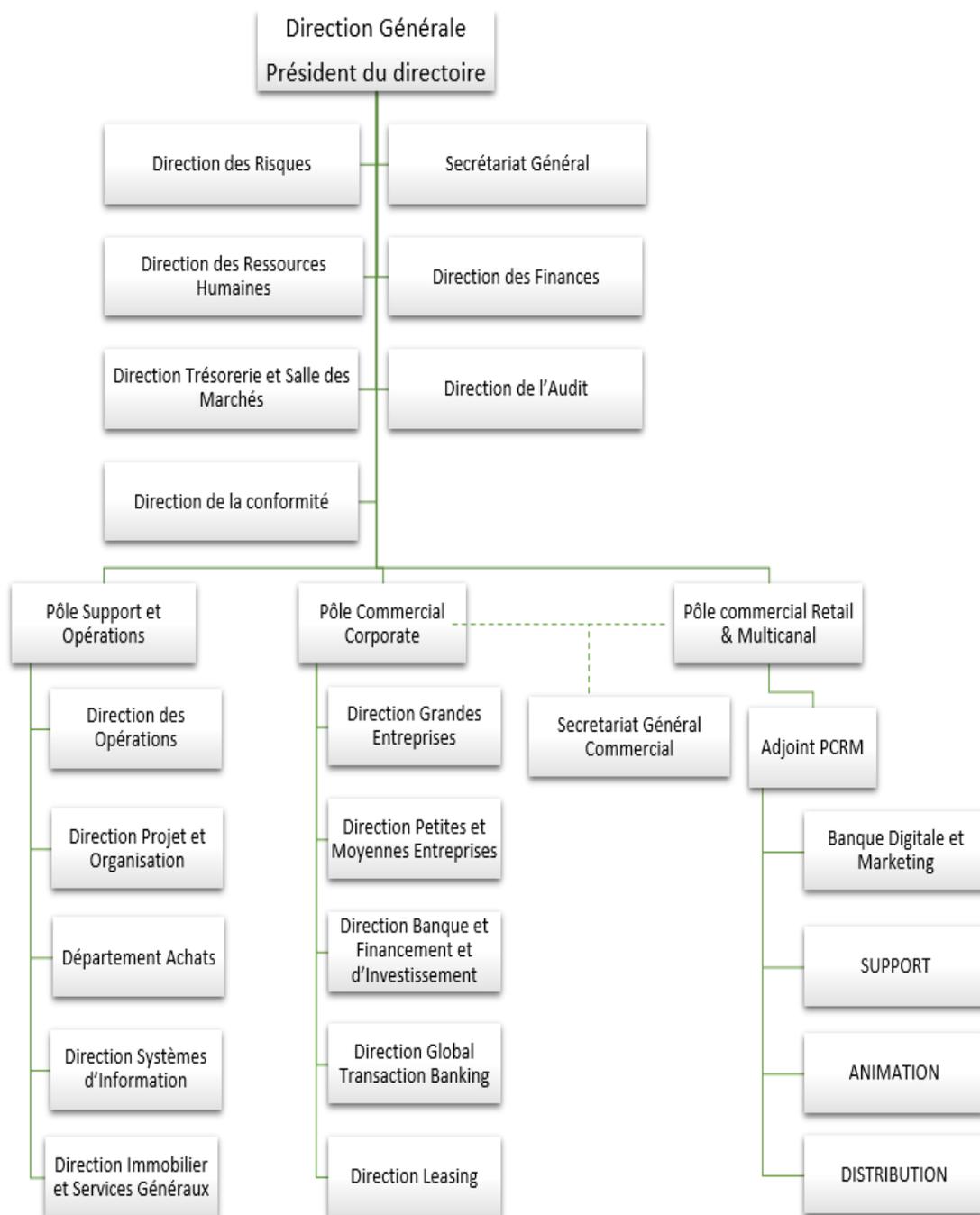


Figure 1: Organigramme de Société Générale Algérie

2.2. Activités de SGA

2.2.1 Activité auprès des entreprises

Depuis plusieurs années, Société Générale Algérie contribue, de manière significative, au financement de l'économie et au développement industriel de l'Algérie. Cela a été particulièrement vrai en 2016, qui a été une année marquée par les mutations visant à encourager le secteur productif. La stratégie de développement commercial de SGA est basée sur son dispositif commercial consacré à la clientèle entreprise qui inclut sa Direction des grandes entreprises, son agence grandes entreprises et ses 11 centres d'affaires dédiés aux pme. [2]

2.2.2 Global Transaction Banking (GTB)

Dans un contexte marqué par la volonté des pouvoirs publics de réduire le niveau des importations, la stratégie de la SGA a été orientée vers l'accompagnement de clients industriels avec des contrats à moyen terme pour favoriser la production. Ce qui a permis de développer une réelle résilience face aux différents changements réglementaires. [2]

2.2.3 Le Marché des Particuliers

Le retour du crédit à la consommation aura permis à Société Générale Algérie de se positionner comme un acteur majeur, citoyen et responsable et ce sont plus de 4000 financements qui ont été mis en place afin d'accompagner la production nationale et le consommateur algérien. Plus de 30 partenariats ont été signés avec des entreprises publiques et privées, plus globalement son offre de produits et services s'est enrichie d'une approche dédiée aux salariés de la fonction publique (Santé et Éducation nationale) auxquels une organisation et un accueil spécifique ont été consacrés en agence. [2]

2.2.4 Le Marché des Professionnels

En 2016 Société Générale Algérie a consolidé sa position sur le marché des professionnels avec une action commerciale très soutenue en direction de ses cœurs de cibles (pharmaciens et professions libérales médicales, tpe) et sa présence à de nombreux Forums dédiés ainsi que l'organisation de journées d'information en direction des entreprises au niveau des chambres de commerce et de l'industrie. Société Générale Algérie a lancé sur ce segment en 2016 la commercialisation des produits de bancassurance, notamment la prévoyance individuelle professionnelle. [2]

2.3. Présentation de la DSI

La Direction des Systèmes d'Information est en charge de la gestion et du pilotage du patrimoine applicatif et technique de la SGA.

Elle est rattachée au Responsable du Pôle Support et Opérations.

2.3.1 Organigramme

C'est au sein du Département Production qui se situe au niveau de la Direction Système d'Information du Pôle Support et Opérations que nous avons effectué notre stage.

Le but principal de notre travail sera d'implémenter une solution de supervision pour permettre à SGA d'avoir une vue globale de ses bases de données Informix, de surveiller leurs performances et surtout pouvoir visualiser différentes métriques en temps réel sous forme de graphiques ergonomiques ce qui permettra de détecter voire d'anticiper des anomalies ou des pannes et d'alerter en cas de dysfonctionnement.

La figure ci-dessous représente l'organigramme de la Direction des Systèmes d'Information :

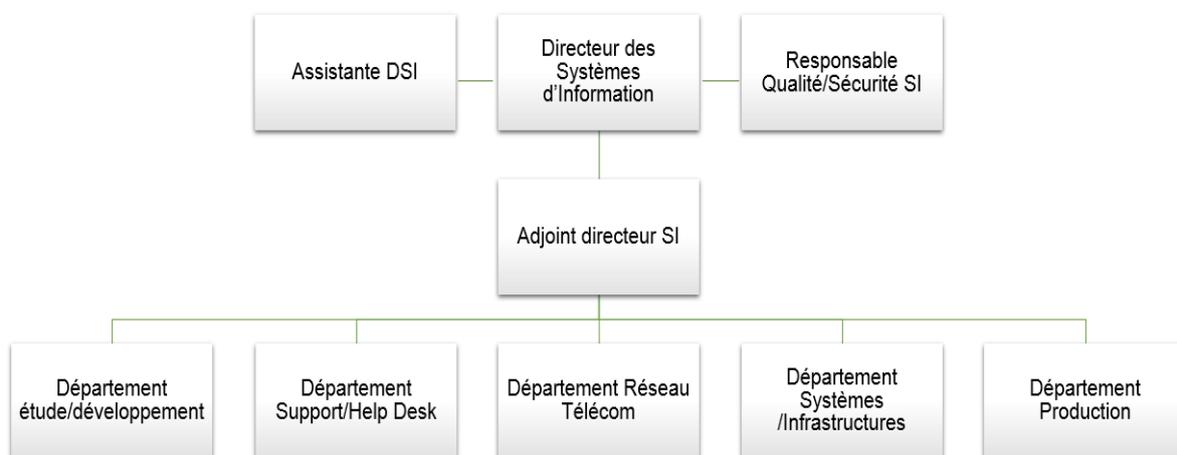


Figure 2: Organigramme de la Direction des Systèmes d'Information

2.3.2 Missions

Les principales missions de la DSI sont les suivantes :

- Assurer les fonctions de développement et de maintenance du patrimoine applicatif de la filiale, dans le respect des normes d'urbanisation et d'architecture définies par le groupe SG.
- Gérer et maintenir en l'état le patrimoine technique téléphonique et bureautique de la filiale dans le respect des normes et standards du groupe SG.
- Accompagner l'ensemble des collaborateurs de la filiale dans la mise en place et l'utilisation des outils logiciels et matériels mis à leurs dispositions.
- Assurer un reporting global et détaillé sur les actions entreprises dans le cadre de son activité.
- S'assurer du respect des normes et règlements internes et de place sur l'utilisation et le fonctionnement du SI.
- S'assurer de la capacité du SI à accompagner la filiale dans le développement de son activité et sur ses différents projets.
- Gérer le budget informatique mis à sa disposition et en assurer un suivi auprès de son entité de rattachement et auprès des entités support du groupe SG.

3. Généralités sur les bases de données

3.1. Définition

Ensemble de données structurées, généralement en champs, organisé en vue de son utilisation par des programmes correspondant à des applications distinctes. Ce regroupement structuré de données, géré par un système de gestion de base de données (SGBD), se réalise de manière à faciliter l'évolution indépendante des données et des programmes. [3]



Figure 3: Base de données

3.2. Objectifs

Les objectifs que l'on assigne généralement aux bases de données et aux systèmes qui les supportent sont les suivants :

- **Exhaustivité** : implique que l'on dispose de toutes les informations relatives au sujet donné.
- **Eviter la redondance et les incohérences des données** : implique l'unicité des informations dans la base de données.
- En général on essaie d'éviter la duplication des données car cela pose des problèmes de cohérence lors des mises à jour de ces données.
- **Offrir un langage de haut niveau** pour la définition et la manipulation des données. Contrôler l'intégrité entre plusieurs utilisateurs et la confidentialité des données. [4]

3.3. Les systèmes de gestion de bases de données (SGBD)

3.3.1 Définition

Un Système de Gestion de Base de Données peut être défini comme un ensemble de logiciels prenant en charge la structuration, le stockage, la mise à jour et la maintenance des données. Autrement dit, il permet de décrire, modifier, interroger et administrer les données. C'est en fait, l'interface entre la base de données et les utilisateurs. [5]

3.3.2 Objectifs

Les mécanismes du système de gestion de base de données visent à assurer :

- **Indépendance physique** : un SGBD offre la facilité de changer le schéma interne sans changer le programme d'application. Cela signifie que le niveau physique peut être modifié indépendamment du niveau conceptuel.
- **Indépendance logique** : le SGBD permet de modifier schéma conceptuel sans changer le programme d'application. Donc le niveau conceptuel peut être modifié sans remettre en cause le niveau physique.
- **La manipulation de données par des langages non procéduraux** : le SGBD doit permettre l'interrogation et la mise à jour de données par des langages de haut niveau spécifiant les données que l'on veut traiter (de quoi) et non pas comment y accéder.
- **La facilité d'administration** : les langages de haut niveau référencent des descriptions logiques des données (schéma externes) stockées dans le dictionnaire de données pour permettre la création et modification de la description.
- **Fiabilité des données** : le SGBD permet de vérifier les contraintes des données ; gérer des transactions (atomicité des transactions) et sécurité (mot de passe, etc.) ; récupérer des données en cas de crash logiciel, OS (Operating System) ou disque. [4]

3.3.3 Fonctionnalités

Un SGBD assure plusieurs fonctionnalités qui sont :

- ✓ La définition et la mise à jour de la structure de la base de données (tables, attributs, vues, index, ...).
- ✓ La recherche des informations utiles en interrogeant la base de données.
- ✓ La manipulation des données de la base et la mise à jour.
- ✓ La définition des droits d'accès pour les différents utilisateurs de la base de données, gestion de la sécurité de la base, confirmation et annulation des transactions. [6]

3.3.4 Conception d'une base de données

Pour simplifier la vision des utilisateurs, les SGBD assurent une abstraction des données stockées sur disque en distinguant trois niveaux de description de données qui sont : [7]

Niveau externe

Au niveau externe, chaque groupe de travail utilisant des données possède une description des données perçues appelée vue ou schéma externe. Cette description est effectuée selon la manière dont le groupe voit la base dans ses programmes d'application. Alors qu'au niveau conceptuel et interne les schémas décrivent toute une base de données, au niveau externe ils décrivent simplement la partie de données présentant un intérêt pour un utilisateur ou un groupe d'utilisateurs. En conséquence, le schéma externe est souvent qualifié de vue externe. Le modèle externe utilisé est dépendant du langage de manipulation de la base de données.

Niveau conceptuel

C'est le niveau central correspondant à la structure canonique de données qui existent dans l'entreprise autrement dit leur structure sémantique inerrance sans souci d'implémentation en machine, représentant la vue intégrée de tous les utilisateurs. La définition du schéma conceptuel d'une entreprise nécessite un accord sur les concepts de base que modélisent les données.

Niveau interne

Ce niveau correspond à la structure de stockage du schéma interne nécessite au préalable le choix d'un SGBD. Elle permet donc de décrire les données telles qu'elles sont stockées dans la machine.

3.3.5 Les systèmes de gestion de bases de données relationnelles

3.3.5.1 Définition d'une base de données relationnelle

Une base de données relationnelle est un ensemble collectif de plusieurs ensembles de données organisés par tables, enregistrements et colonnes. Les bases de données relationnelles établissent une relation bien définie entre les tables. [8]

3.3.5.2 Définition d'un SGBDR

Un SGBDR (système de gestion de base de données relationnelle) est un système (logiciel) qui permet de gérer une base de données relationnelle. Les SGBDRs peuvent être soit de type client/serveur (Oracle, SQL Server, MySQL, PostgreSQL, etc.), soit de type fichiers partagés (Access, SQL Server CE, Paradox, dBase, etc.).[4]

3.3.5.3 Propriétés

- Base formelle reposant sur des principes parfaitement définis.
- Organisation structurée des données dans des tables interconnectées (d'où le qualificatif relationnelles), pour pouvoir détecter les dépendances et redondances des informations.
- Implémentation d'un langage relationnel ensembliste permettant à l'utilisateur de décrire aisément les interrogations et manipulation qu'il souhaite effectuer sur les données.
- Indépendance des données vis-à-vis des programmes applicatifs (dissociation entre la partie "stockage de données" et la partie "gestion" ou "manipulation").
- Gestion des opérations concurrentes pour permettre un accès multi-utilisateur sans conflit.

- Gestion de l'intégrité des données, de leur protection contre les pannes et les accès illicites. [9]

3.4. Classification des bases de données

3.4.1 Classification selon l'usage

✓ Base de données opérationnelles OLTP

Les bases de données dites opérationnelles ou OLTP (OnLine Transaction Processing) sont destinées à assister des usagers à tenir l'état d'activités quotidiennes. Elles permettent en particulier de stocker sur le champ les informations relatives à chaque opération effectuée dans le cadre de l'activité : achats, ventes, réservations, paiements. Dans de telles applications l'accent est mis sur la vitesse de réponse et la capacité de traiter plusieurs opérations simultanément.

✓ Base de données d'analyse OLAP

Les bases de données d'analyse dites aussi OLAP (OnLine Analytical Processing), sont composées d'informations historiques telles que des mesures sur lesquelles sont effectuées des opérations massives en vue d'obtenir des statistiques et des prévisions. Les bases de données sont souvent des entrepôts de données (anglais datawarehouse) : des bases de données utilisées pour collecter des énormes quantités de données historiques de manière quotidienne depuis une base de données opérationnelle.

Le contenu de la base de données est utilisé pour effectuer des analyses d'évolution temporelle et des statistiques telles que celles utilisées en management.

Dans de telles applications l'accent est mis sur la capacité d'effectuer des traitements très complexes. [4]

3.4.2 Classification selon l'emplacement

✓ Bases de données centralisées

Une base de données centralisée est une base de données gérée par un seul SGBD et stockée dans sa totalité à un emplacement physique unique. Ses divers traitements sont confiés à une seule unité de traitement. [4]

✓ Bases de données réparties

Une base de données répartie (BDR) est un ensemble structuré et cohérent de données, stocké sur des processeurs distincts, et géré par un système de gestion de bases de données réparties (SGBDR). [10]

4. Base de données Informix**4.1. IBM Informix Database Server**

IBM Informix est un serveur de base de données rapide et évolutif qui permet de gérer des bases de données relationnelles, de relations aux objets et dimensionnelles.

Le serveur de base de données IBM Informix s'exécute sur les systèmes d'exploitation UNIX, Linux, Mac OS X et Windows. [11]

4.2. IBM Informix database

Une base de données intégrée et sécurisée, optimisée pour les données OLTP et Internet of Things (IoT). Informix a la capacité unique d'intégrer de manière transparente SQL, NoSQL / JSON, des séries chronologiques et des données spatiales. [11]

4.3. Architecture du serveur de base de données Informix

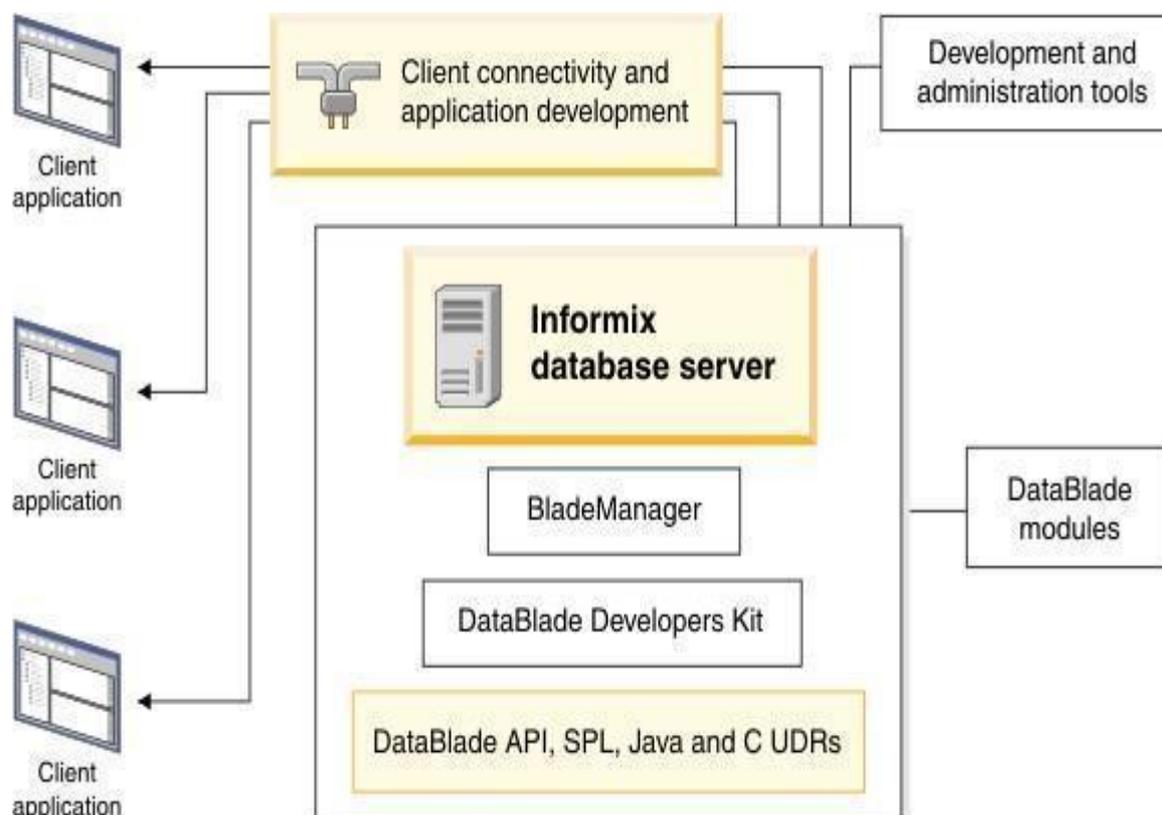


Figure 4: Architecture du serveur de base de données Informix

Toutes les éditions d'Informix contiennent les outils client suivants en plus du serveur de base de données : [11]

IBM Informix Client Software Development Kit (Client SDK) et IBM Informix Connect :

Client SDK Contient des API client pour le développement et l'exécution d'applications client et contient l'outil OAT (OpenOffice) IBM pour Informix pour la surveillance et la gestion du serveur de base de données.

IBM Informix Connect contient uniquement les bibliothèques d'exécution des API client pour permettre aux applications qui s'exécutent sur les ordinateurs clients d'accéder au serveur de base de données.

DataBlade Developers Kit (DBDK)

Contient des outils permettant de développer et de regrouper les modules DataBlade, qui sont des progiciels qui étendent les fonctionnalités du serveur de base de données.

IBM Informix Blade Manager

Un utilitaire pour inscrire et désinscrire les modules DataBlade.

4.4. Informix vs Oracle

Plusieurs SGBDs sont disponible sur le marché, nous nous intéressons à faire une comparaison entre Informix et Oracle comme le tableau ci-dessous le montre : [12][13]

	IBM Informix	Oracle
Développeur	IBM	Oracle Corporation
Système d'exploitation	AIX HP-UX Linux Solaris Windows OS X	AIX HP-UX Linux Solaris Windows z/OS OS X
Version	IDS IDS Enterprise IDS Workgroup IDS Express IDS Developer	Oracle Enterprise Edition Oracle Standard Edition Oracle Personal Edition Oracle Express Edition
Scripts côté serveur	SPL	PL/SQL
Avantages	<ul style="list-style-type: none"> ➤ Solution très stable. ➤ Facilité d'administration. ➤ Informix dispose d'un système de réplication très riche et étendu 	<ul style="list-style-type: none"> ➤ Richesse fonctionnelle. ➤ Fonction d'audit évolué. ➤ Row Level Storage Security (RLSS). ➤ Intégration LDAP, SSL, Unicode.

	<p>offrant diverses fonctionnalités utiles.</p> <ul style="list-style-type: none"> ➤ Solution de haute disponibilité. ➤ Performances exceptionnelles du Data Warehouse grâce au module Informix Warehouse Accelerator. ➤ Fournit la programmation compatible avec MongoDB et supporte les APIs MongoDB. ➤ Son moteur hybride permet une vraie intégration du monde non structuré (NoSQL) avec le monde structuré (SQL). ➤ Les jointures sont permises entre collections JSON et entre collections JSON et tables SQL 	<ul style="list-style-type: none"> ➤ Réplication intégrée. ➤ Haute disponibilité. ➤ Assistants performants via Oracle Manager Server. ➤ Possibilité de gérer en interne des tâches et des alarmes. ➤ Pérennité de l'éditeur. ➤ Accès aux données système via des vues, bien plus aisément manipulable que des procédures stockées. ➤ Interface utilisateur remaniée et extrêmement riche. ➤ Services Web, support XML. ➤ Compression des données et des sauvegardes.
Inconvénients	<ul style="list-style-type: none"> ➤ Le nom d'Informix est presque invisible sur le marché. ➤ La politique de prix est très complexe. ➤ Pérennité de la solution. ➤ Très peu d'outils graphiques disponible pour l'administration. 	<ul style="list-style-type: none"> ➤ Prix élevé, tant au point de vue des licences que des composants matériels (RAM, CPU) à fournir pour de bonnes performances. ➤ Administration complexe, liée à la richesse fonctionnelle. ➤ Fort demandeur de ressources, Oracle est bien plus gourmand en ressource mémoire que ses concurrents.

Tableau 1: Comparaison entre Informix et Oracle

4.5. Caractéristiques

- Les fonctionnalités objets d'Informix IDS permettent de mixer les données Time Series avec des tables traditionnelles dans le même schéma de base en bénéficiant de tous les avantages d'Informix IDS : backup en ligne, réplication en cluster, réplication grande distance, grid, compression, données JSON, accélérateur de requête décisionnelle, fiabilité et simplicité d'administration. [14]
- Informix est hybride : Informix comprend SQL et NoSQL (99 +% compatible avec MongoDB) : il permet même d'exécuter des requêtes de style SQL sur des collections NoSql et / ou d'interroger des données SQL en utilisant la syntaxe MongoDB. Il permet même d'exécuter des jointures entre les données NoSql et SQL dans une même requête. [15]
- L'hétérogénéité des données à traiter constitue un défi de taille, que ce soit en termes de matériel, en termes applicatifs et enfin de stockage des données. Trois défis auxquels IBM a décidé de répondre avec Informix. Avec NoSQL et JSON/BSON en type de données NATIF, Informix répond parfaitement au caractère excessivement hétérogène des données à collecter. [16]

4.6. Informix et le Cloud

BlueMix est une plateforme IBM de développement sur le Cloud. Une solution originale très intéressante dans son concept, permettant aux développeurs de développer et tester leurs applications Nodes.js, Java et autres systèmes de développement du monde libre, Informix faisant partie des modules proposés, sous le nom de Time Séries. Cette plateforme de développement intégrée est tout particulièrement adaptée au développement d'applications dans le monde des Objets connectés (Internet of Things) [16]

5. Conclusion

A travers ce chapitre nous avons présenté l'organisme d'accueil Société Générale Algérie plus précisément la Direction des Systèmes d'Information et les différentes missions qu'elle assure, nous avons ensuite introduit quelques généralités sur les bases de données en général notamment les bases de données Informix.

Chapitre II : La supervision de bases de données

1. Introduction

L'informatique est devenue un élément indispensable pour une entreprise quelque soit son secteur d'activité. Les problèmes liés à l'Informatique doivent être réduits au maximum, car une indisponibilité du système d'information aura des impacts préjudiciables sur l'activité et la notoriété de l'entreprise. Dans ce chapitre nous allons aborder la notion de supervision informatique en général et plus particulièrement celle des bases de données.

2. La supervision

2.1. Définition

La supervision, dénommée aussi **surveillance informatique** ou **monitoring**, est le procédé qui consiste à Contrôler et Vérifier, l'activité d'un système informatique.

Les principales solutions de supervision s'intègrent dans l'activité quotidienne informatique et permettent de surveiller, analyser, piloter en agissant directement après avoir reçu des alertes informant de potentielles anomalies.

Ces alertes peuvent être remontées via des scripts, des mails, des fax, des appels vocaux ou bien par l'envoi de simple SMS d'alerte. [17]

2.2. Principe

Les systèmes d'information étant de plus en plus complexes, leur surveillance et la localisation des problèmes deviennent de plus en plus ardues pour l'administrateur réseaux et systèmes.

Le service informatique ainsi que l'administrateur doivent connaître à tout moment l'état de chaque équipement et service sur le réseau pour une plus grande réactivité. [18]

2.3. Types

On distingue trois types de supervision :

✓ **Supervision système**

La supervision système porte principalement sur les trois principales ressources systèmes, c'est-à-dire l'utilisation du processeur, de la mémoire (RAM) et de l'espace disque.

✓ **Supervision réseau**

La supervision réseau porte sur la surveillance de manière continue de la disponibilité des services en ligne.

✓ **Supervision applicative**

La supervision des applications (ou supervision applicative) permet de connaître la disponibilité des machines en termes de services rendus en testant les applications hébergées par les serveurs.

Plus concrètement la supervision permet de :

- Prévenir en cas d'incident par mail, sms, etc.
- Détecter des anomalies de sécurité ou du système.
- Surveiller l'état des différents périphériques.
- Surveiller les différentes applications.
- Surveiller l'état des différents systèmes d'information.
- Alerter en cas d'interruption d'un service ou d'une application.
- Garantir et assurer la disponibilité des services.
- Informer sur l'utilisation de la ressource et des performances du système.
- Interroger diverses sondes permettant de remonter des informations. [18]

2.4. Aspects

La supervision peut porter sur plusieurs aspects de l'informatique.

Pour simplifier on peut les classer dans trois catégories principales qui sont les suivantes :

- **La Fiabilité** : Il s'agit de l'utilisation la plus courante du monitoring informatique. Une surveillance permanente de la disponibilité de l'équipement est effectuée, et ce pour détecter la moindre anomalie et de la signaler à l'administrateur.
- **La Performance** : Le monitoring de performance a pour but de retourner des informations sur le rendement d'un équipement ou d'un service comme par exemple le temps de résolution DNS, le temps de connexion, le temps de récupération du premier octet et dans le cas d'une page Web, le temps de récupération de la page et de l'ensemble des éléments de celle-ci (image, scripts...).
- **Le Contenu** : Dans ce dernier cas, on analyse les informations retournées par les éléments surveillés pour détecter, par exemple, la suppression d'un fichier sur un serveur FTP, la modification d'une page Web ou encore la disparition d'un mot clef.

Toutes ces approches sont complémentaires et peuvent bien entendu se retrouver regroupées dans une seule analyse. L'analyse Fiabilité étant implicitement présentée dans le cas d'une analyse Performance et Contenu. [19]

2.5. Quelques Standards normalisés

Les systèmes de supervision utilisent des protocoles, très réglementés par la **DMTF (Distributed Management Task Force)** depuis 2005, parmi les principaux utilisés :

- **La supervision des protocoles IPMI (Intelligent Platform Management Interface)** :
C'est l'un des plus utilisés, il concerne surtout les serveurs et cette interface intelligente de gestion de matériel permet, entre autres, de contrôler à distance certains composants très sensibles comme les sondes et autres ventilateurs.
- **La supervision des protocoles JMX (Java Management Interface)** :

C'est l'API qui permet de gérer une application en cours d'exécution. JMX est maintenant complètement intégré dans J2SE à partir de la version V.

➤ **La supervision protocole SNMP** (Simple Network Management Protocol) :

C'est le protocole de communication et de gestion simplifiée du réseau. C'est le **SNMP** qui permet aux administrateurs de contrôler et de gérer (diagnostiquer) tous les éléments actifs du réseau. En langage SNMP on ne supervise pas, on manage, mais le résultat est similaire. Il est composé de 3 éléments essentiels : le superviseur, les nodes (ou nœuds en français) et les agents.

➤ **La supervision protocole ITIL** (Information Technology Infrastructure Library) :

C'est une norme, ensemble de bonnes pratiques pour la bonne gestion d'un système d'information. Né en Grande Bretagne, et populaire en Europe depuis plus de 35 ans, il tend à s'implanter aux USA grâce à l'impulsion de certaines grandes SSII.

Basé sur la CMDB (Configuration Management DataBase) : c'est une BDD qui unifie les composants d'un système d'information et qui en plus, permet de comprendre l'organisation et de modifier la configuration si nécessaire. C'est la base même du système **ITIL** en positionnant des blocs organisationnels et des flux d'informations, les recommandations ITIL, abordent des sujets aussi variés que :

- Les productions informatiques.
- Les réductions des risques.
- L'augmentation de la qualité du service.
- L'efficacité globale du système d'information.

- **La supervision protocole SBLIM** (Standard Based Linux Instrumentation for Manageability) :

SBLIM, est nommé par les experts en langage courant SUBLIME. Il s'applique aux machines LINUX et permet entre autres d'avoir accès aux technologies WBEM (Web Based Enterprise Management). Ce standard est exclusivement mis en avant par IBM qui en assure aussi le développement.

- **La supervision protocole WBEM** (Web Based Enterprise Management) :

WBEM est un ensemble de standards de base intégrés dans les outils de supervision, pour faciliter l'échange entre plateformes et technologies. WBEM sont des standards Internet de gestion, surtout développés pour unifier les environnements dans l'informatique distribuée.

- **La supervision protocole WS-MANAGEMENT** (Web Services for Management) :

Fournit la méthodologie pour échanger des informations d'administration à travers les infrastructures IT, spécification fournie par le DTMF. Basé sur les Web Services (SOAP), il est très proche du protocole WBEM.

- **La supervision protocole Windows Management Instrumentation WMI**

C'est le protocole pour les plateformes Windows, il étend le modèle du C I M, pour représenter les objets, dans cet environnement. Son interface cohérente et orientée objet, utilise les normes de l'industrie et permet aux informaticiens une utilisation simplifiée des tâches de gestion. L'accès aux données sous WMI, que ce soit en local ou à distance, est totalement transparent.

Tous ces protocoles sont normalisés et gérés par la DMTF qui est un organisme mondial, où sont regroupés tous les grands constructeurs et donneurs d'ordre, qui met en place et gère les standards de la technologie.

La **DMTF** simplifie la gérabilité de tous ces standards, grâce à la collaboration et à la participation des grandes sociétés mondiales de technologie, ainsi que des principaux constructeurs. [17]

2.6. Intérêt

La supervision informatique permet de surveiller, d'analyser, d'alerter et d'agir sur les fonctionnements normaux et anormaux des différents systèmes informatiques en production. Plus précisément, elle consiste à surveiller l'état de l'ensemble de l'infrastructure réseau, des serveurs et du système d'information permettant d'anticiper les différents incidents et ainsi de diagnostiquer rapidement une panne. [18]

3. Supervision de bases de données

3.1. Définition

La surveillance de base de données consiste à évaluer la disponibilité, les performances et les fonctionnalités de la base de données. [20]

3.2. Métriques de supervision

On distingue deux types de métriques : [21]

a) Work metrics

Les métriques de travail indiquent l'état de santé d'un système en mesurant son rendement utile, on distingue :

- **Le débit** est le nombre d'unités d'information qu'un système peut traiter par unité de temps. Le débit est généralement enregistré en tant que nombre absolu.
- **Les métriques de réussite** représentent le pourcentage de travail exécuté avec succès.
- **Les métriques d'erreur** capturent le nombre de résultats erronés, généralement exprimés en taux d'erreurs par unité de temps.
- **Les métriques de performance** quantifient l'efficacité avec laquelle un composant effectue son travail. La métrique de performance la plus courante est **la latence**, qui représente le temps nécessaire pour compléter une unité de travail.

b) Métriques de ressources

La plupart des composants d'une infrastructure logicielle servent de ressources à d'autres systèmes.

Certaines ressources sont de bas niveau : par exemple, les ressources d'un serveur incluent des composants physiques tels que **le processeur, la mémoire, les disques et les interfaces réseaux**.

Mais un composant de plus haut niveau, tel qu'une base de données ou un microservice de géolocalisation peut également être considéré comme une ressource si un autre système a besoin de ce composant pour effectuer son travail.

Les métriques de ressources peuvent aider à reconstituer une image détaillée de l'état d'un système.

Pour chaque ressource on peut collecter des métriques couvrant quatre domaines clés :

- **L'utilisation** est le pourcentage de temps pendant lequel la ressource est occupée, ou le pourcentage de la capacité de la ressource en utilisation.
- **La saturation** est une mesure de la quantité de travail demandé que la ressource ne peut pas encore traiter, souvent en file d'attente.
- **Les erreurs** représentent des erreurs internes qui peuvent ne pas être observables dans le travail produit par la ressource.
- **La disponibilité** est le pourcentage de temps durant lequel la ressource est capable de remplir sa fonction prévue.

3.3. Surveillance des performances de bases de données

Le contrôle des performances de la base de données est un effort actif pour trouver des problèmes et les résoudre, ainsi que pour trouver des moyens d'améliorer les performances de la base de données. Grâce à divers outils et interfaces, un administrateur de base de données peut surveiller les performances de diverses parties de la base de données desservant différentes applications et obtenir une visibilité sur les performances de la base de données via une interface utilisateur spécialisée ou Web. [22]

3.3.1 Objectifs

L'objectif principal de la surveillance des performances de la base de données est d'évaluer le fonctionnement d'un serveur de base de données, à la fois matériel et logiciel. Cela implique de prendre des instantanés réguliers des indicateurs de performance au fil du temps afin de déterminer l'heure exacte à laquelle des problèmes tels que les goulots d'étranglement se produisent afin de déterminer exactement ce qui a causé les problèmes à ce moment précis et trouver une bonne solution. [22]

C'est parce que la plupart des problèmes se produisent à des moments différents, et il n'est tout simplement pas possible pour un administrateur de surveiller la performance en temps réel tout le temps. Donc, regarder les données historiques est généralement l'approche de la surveillance. La plupart des outils de surveillance ont une sorte de système d'alarme et de notification pour informer l'administrateur des problèmes en cours.

3.3.2 Avantages

Les avantages de la surveillance des performances de la base de données sont les suivants : [22]

- Déterminer si les performances peuvent être améliorées.
- Déterminer les failles de performances et de sécurité en évaluant l'activité de l'application et de l'utilisateur.

3.4. Type d'outils de surveillance des bases de données

1. **Les outils de surveillance Serveur / Infrastructure** : recueillent souvent les métriques émises par les systèmes de bases de données via le système d'exploitation. Par exemple, Nagios ou Microsoft Operations Manager, collectent des métriques du système d'exploitation et effectuent des appels basiques pour collecter des métriques et des journaux. Ceux-ci seront utilisés pour indiquer la santé de la base de données en général, mais ils ne parviennent pas à approfondir.

2. Les outils de surveillance des performances des bases de données

Ces outils vont profondément dans la base de données en cours pour comprendre les performances opérationnelles. Comment fonctionnent les index, quelles sont les principales requêtes SQL, quels sont les travaux en cours d'exécution, quels systèmes effectuent le plus d'appels à la base de données et beaucoup d'autres données pertinentes pour un DBA. (ex: Oracle Enterprise Manager, Dell Foglight, Idera Diagnostics Manager, AppDynamics pour les bases de données).

3. Les outils de surveillance des performances applicatives

Surveillent les applications exécutées depuis le client (navigateur, mobile, etc.) via les serveurs d'applications, dans les appels de base de données et les performances de bout en bout des transactions (ex : AppDynamics, Dynatrace et autres). [23]

3.5. Quelques outils de supervision

Nagios

Est un logiciel de supervision réseau libre sous licence GPL qui fonctionne sous Linux. Il a pour fonction de surveiller les hôtes et services spécifiés, alertant l'administrateur des états des machines et équipements présents sur le réseau.

Bien qu'il fonctionne dans un environnement Linux, ce logiciel est capable de superviser toutes sortes de systèmes d'exploitation (Windows XP, Windows 2000, Windows 2003 Server, Linux, Mac OS entre autres) et également des équipements réseaux grâce au protocole SNMP.

Cette polyvalence permet d'utiliser Nagios dans toutes sortes d'entreprises, quel que soit la topologie du réseau et les systèmes d'exploitation utilisés au sein de l'entreprise.

Ce logiciel est composé de trois parties :

- Le moteur de l'application qui gère l'ordonnance et les supervisions des différents équipements.

- Les Plugins qui servent d'intermédiaire entre les ressources que l'on souhaite superviser et le moteur de Nagios. Il faut bien noter que pour accéder à une certaine ressource sur un hôte, il faut un plugin coté Nagios et un autre coté hôte administré.

-L'interface web qui permet d'avoir une vue d'ensemble des états de chaque machine du parc informatique supervisé et ainsi pouvoir intervenir le plus rapidement possible en ciblant la bonne panne. [23]

Fonctionnalités

Nagios dispose de plusieurs fonctionnalités nous pouvons citer : [24]

- La supervision des services réseaux (SMTP, http...), des hôtes et des ressources systèmes (CPU, charge mémoire...).
- Représentation colorée des états des services et des hôtes.
- Génération de rapports.
- Cartographie du réseau.
- Gestion des alertes.
- Surveillance des processus (sous Windows, Unix...).
- La supervision à distance peut utiliser SSH ou un tunnel SSL.
- Les plugins sont écrits dans les langages de programmation les plus adaptés à leur tâche (Bash, C++, Python, Perl, PHP, C#, etc.)



Figure 5: Fonctionnalités de Nagios

Avantages

- Solution complète permettant le reporting, la gestion des pannes et d'alarmes, gestion des utilisateurs.
- Pléthore de plugins qui permettent d'étendre les possibilités de Nagios.

Inconvénients

Les principaux inconvénients de Nagios sont :

- Configuration complexe mais peut s'améliorer en ajoutant Centreon.
- Interface peu ergonomique et intuitive.

Centreon

Centreon est un logiciel libre de supervision Open Source, édité par la société française Centreon. Il mesure la disponibilité et la performance des couches applicatives, du service utilisateur jusqu'aux ressources matérielles. Centreon est un puissant outil destiné aux administrateurs et exploitants du service de supervision.

Il intègre depuis 2012 son propre moteur de collecte (Centreon Engine) et gestionnaire d'événements (Centreon Broker).

Au-delà des concepts de configuration qui sont les plus évolués dans l'écosystème Nagios, Centreon apporte de nombreuses fonctionnalités telles que la consultation de l'état des services et des machines supervisés, la métrologie, le reporting, l'accès aux événements de supervision, la gestion avancée des utilisateurs via des listes de contrôle d'accès (ACL), etc. [25]

Hosts	Services	Status	Duration	Last Check	Tries	Status information
pj-front-is	CPU	UNKNOWN	24m 47s	22/01/2010 17:05:11	1/3 (H)	ERROR: when getting SNMP version : No response from remote host 'pj-front-is'.
	Disk-C	UNKNOWN	24m 17s	22/01/2010 17:05:11	1/3 (H)	ERROR: hrStorageDescr Table : No response from remote host 'pj-front-is'.
	Disk-D	UNKNOWN	27m 39s	22/01/2010 17:05:11	1/3 (H)	ERROR: hrStorageDescr Table : No response from remote host 'pj-front-is'.
	Disk-E	UNKNOWN	24m 13s	22/01/2010 17:05:11	1/3 (H)	ERROR: hrStorageDescr Table : No response from remote host 'pj-front-is'.
	FTP	CRITICAL	47m 29s	22/01/2010 17:04:24	1/3 (H)	CRITICAL - Socket timeout after 10 seconds
	HTTP	CRITICAL	28m 13s	22/01/2010 17:05:24	1/3 (H)	CRITICAL - Socket timeout after 10 seconds
	Memory	UNKNOWN	27m 47s	22/01/2010 17:05:11	1/3 (H)	ERROR: hrStorageDescr Table : No response from remote host 'pj-front-is'.
	Ping	CRITICAL	28m 13s	22/01/2010 17:05:47	3/3 (H)	GPING CRITICAL - --- pj-front-is ping statistics ---
	Swap	UNKNOWN	28m 13s	22/01/2010 17:05:11	1/3 (S)	ERROR: hrStorageDescr Table : No response from remote host 'pj-front-is'.
	Back MySQL	Disk-/	OK	1d 53m 12s	22/01/2010 17:03:25	1/3 (H)
Back Memcached	Disk-/mnt	OK	1d 52m 34s	22/01/2010 17:04:18	1/3 (H)	Disk OK - /mnt TOTAL: 413.378GB USED: 0.194GB (0%) FREE: 413.182GB (100%)
	Load	OK	1d 51m 57s	22/01/2010 17:04:47	1/3 (H)	load average: 0.00, 0.00, 0.00.
	MySQL	OK	23h 18m 42s	22/01/2010 17:02:57	1/3 (H)	TCP OK - 0.001 second response time on port 3306
	Ping	OK	1d 50m 4s	22/01/2010 17:01:46	1/3 (H)	GPING OK - rtt min/avg/max/mdev = 0.346/0.414/0.487/0.062 ms
Back Memcached	Traffic	OK	1d 52m 40s	22/01/2010 17:03:58	1/3 (H)	Traffic In: 0.00 b/s (0.0 %), Out: 0.00 b/s (0.0 %) - Total RX Bits In: 15.70 GB, Out: 21.13 Gb
	Traffic-Limited	OK	1d 53m 3s	22/01/2010 17:03:52	1/3 (H)	Traffic In: 3.44 kb/s (0.2 %), Out: 43.43 kb/s (2.2 %) - Total RX Bits In: 15.70 GB, Out: 21.13 Gb
Back Memcached	Disk-/	OK	1d 2h 21m 19s	22/01/2010 17:02:47	1/3 (H)	Disk OK - / TOTAL: 9.843GB USED: 2.699GB (27%) FREE: 7.144GB (73%)
	Disk-/mnt	OK	1d 2h 20m 33s	22/01/2010 17:03:24	1/3 (H)	Disk OK - /mnt TOTAL: 413.378GB USED: 0.194GB (0%) FREE: 413.182GB (100%)

Figure 6: Centreon

Avantages

- La robustesse et la renommée de Nagios.
- Une interface beaucoup plus sympathique, permettant de tout configurer, de garder un œil sur tout le réseau en permanence.
- Les utilisateurs de Nagios ne seront pas perdus pour autant, l'interface reprenant avantageusement certaines vues Nagios. [24]

Inconvénients

Centreon possède des inconvénients :

- L'interface peut paraître complexe car il existe beaucoup d'options.
- Un développement qui n'est pas encore en phase avec celui de Nagios : Parfois des problèmes de compatibilité.
- Un peu plus lourd que du Nagios pur. [24]

HP OpenView

Est un outil de supervision reconnu sur le marché. Son principal avantage est la centralisation des informations sur un serveur de supervision. Il a pour rôle de gérer et de surveiller entre autres les infrastructures et les services réseaux. Il est destiné aux moyennes et grandes entreprises qui souhaitent avoir une vue globale de leur réseau et de son état.

HP OpenView dispose d'un outil permettant de superviser **les bases de données**, « HP OpenView – Smart Plug-In for Database ».

The Smart Plug-in (SPI) for Databases ajoute des capacités de monitoring de bases de données à HP OpenView Operations pour Windows. Le plugin Smart Plug-In for Database (SPI) peut surveiller les bases de données **Oracle, Microsoft SQL Server, Sysbase et Informix.**

Ce type de plugin permet de :

- Augmenter la disponibilité et les performances des serveurs de bases de données.
- Réduire les coûts de support associés aux serveurs de bases de données.
- Améliorer la gestion de la capacité et de la manipulation des serveurs de bases de données. [18]

Zabbix

Zabbix est un logiciel open source. Il permet de surveiller le statut de divers services réseaux, l'intégrité des serveurs et autres matériels réseaux. Il utilise un mécanisme de notification flexible qui permet aux utilisateurs de configurer une base d'alerte e-mail pour pratiquement tous les évènements. Cela permet une réponse rapide aux problèmes serveurs. Il offre aussi un excellent reporting et des fonctionnalités de visualisation de données basées sur les données stockées.

Zabbix offre un certain nombre de fonctionnalités :

- Découverte automatique des serveurs et des périphériques réseaux
- Logiciel Multiplateforme Windows, Linux, Unix, AIX, Free BSD, ...etc.
- Supervision sans agent.
- Authentification d'agent sécurisée.
- Permission utilisateurs flexibles.
- Interfaces web.
- Notification par e-mail d'évènement prédéfinis.
- Haut niveaux de visualisation des ressources supervisées.

Zabbix dispose d'un plugin pour pouvoir surveiller **les bases de données Oracle** : «**Orabbix** ». C'est un plugin conçu pour assurer le suivi de plusieurs niveaux, la performance et les rapports de disponibilité et de mesure pour les bases de données Oracle avec des indicateurs de performance du serveur. Il fournit néanmoins un mécanisme efficace d'acquérir des données à partir de nombreuses instances d'Oracle et de fournir cette information au serveur Zabbix.

Il utilise les fonctionnalités de reporting pour toutes les données recueillies et fournit une analyse tels que des graphiques et des mesures du niveau de service pour les parties pertinentes. La répartition actuelle contient un ensemble de modèle prédéfinis qui intègre l'alerte et la représentation graphique des capacités de déploiement. [18]

3.6. Les contraintes de la supervision de bases de données

La supervision de bases de données correspond à une partie du travail d'un administrateur de bases de données. Cette méthode ne remplacera jamais celle qu'exerce un administrateur de bases de données qui surveille constamment ces bases et améliore au fur et à mesure ces méthodes de supervision permettant d'avoir un serveur de base de données performant. La contrainte principale de cette méthode est de savoir quels types d'outils seraient plus adaptés pour prendre en compte les différentes versions de bases de données. [18]

4. Conclusion

A travers ce chapitre nous avons exposé le principe de la supervision informatique ainsi que la supervision des bases de données qui est l'objet de notre travail, et nous avons ensuite présenté quelques outils de supervision tels que Nagios, Centreon, HP OpenView et Zabbix.

Chapitre III : Conception

1. Introduction

Dans le précédent chapitre nous avons présenté le principe de la supervision des bases de données ainsi que les différents outils existants sur le marché, à travers celui-ci nous délimitons le contexte de notre projet, définissons les outils utilisés et présentons les différentes étapes de la démarche suivie pour la mise en place de notre solution de supervision.

2. Contexte du projet

Les bases de données sont le noyau central de tout système d'information, quand ce dernier se trouve au cœur du business de l'entreprise, comme c'est le cas des banques, celles-ci deviennent d'une importance capitale.

Le suivi des indicateurs de performance des bases de données à travers une plateforme de supervision reste le meilleur moyen de détection des anomalies et de la prévention de pannes qui peuvent menacer le bon fonctionnement de l'entreprise.

Pour veiller sur la bonne santé des bases de données, afin de garantir sa disponibilité, et après de nombreux échanges avec l'équipe opérationnelle chargé de l'exploitation des serveurs de bases de données. Nous avons proposé notre projet qui répond au besoin de l'entreprise qui consiste à mettre en place une solution de surveillance en temps réel de ses différentes bases de données afin d'être prévenu en cas d'incidents.

Cette solution va concerner plus précisément les bases de données INFORMIX.

3. Schéma de la base de données supervisée

Afin de superviser notre système de base de données, nous procédons à la collecte des métriques depuis la base de données SYSMASTER, et plus précisément les tables *system-monitoring interface* (SMI), ces tables contiennent les informations sur l'état, l'utilisation des ressources et l'activité des sessions du server de base de données.

Pour simuler l'activité de notre serveur de base de données, nous aurons besoin d'un schéma de la base de données à requêter de manière similaire qu'en production, cela en envoyant des milliers de requêtes par minutes.

Dans le cas de notre étude, et pour des raisons de confidentialité des schémas des bases de données de Société Générale Algérie, nous avons choisi de reprendre un exemple très connu de schéma de base de données, le HR Schéma d'Oracle. Celui-ci reprend les tables d'une base de données d'un système de gestion des ressources humaines. On va procéder à la création des tables dans une base de données Informix, et intégrer les données exemples.

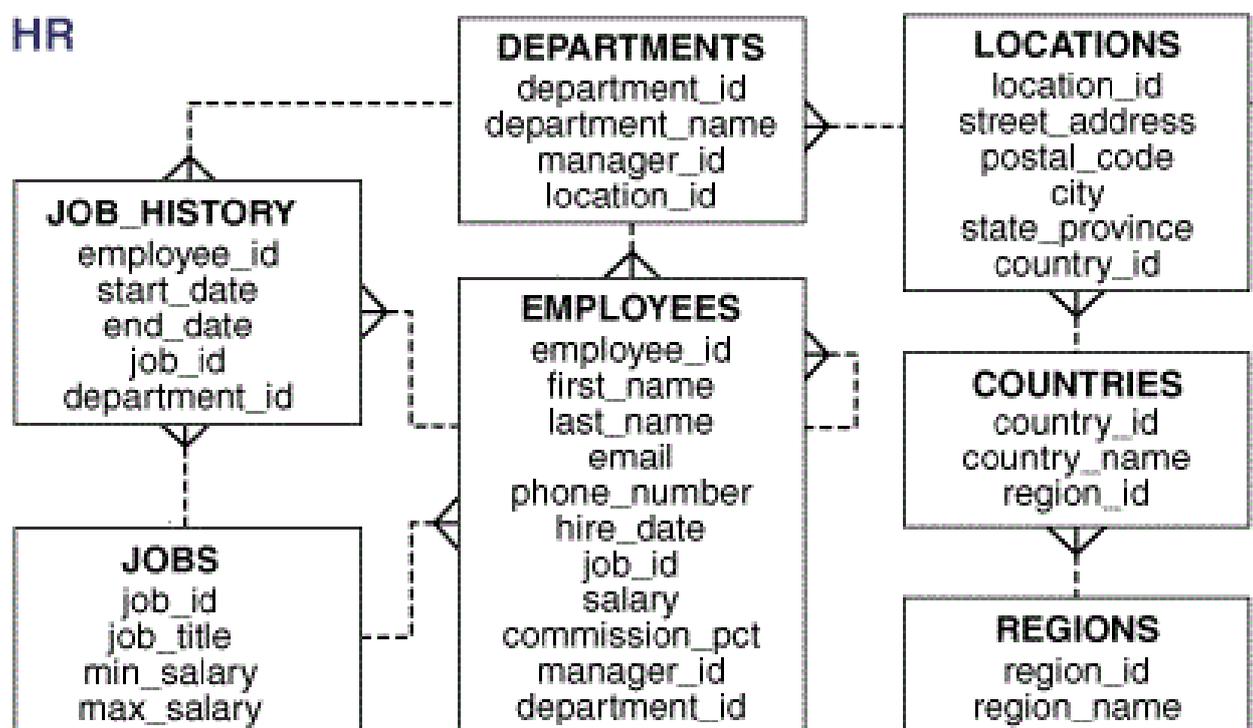


Figure 7: Schéma de base de données

4. Architecture de la solution

La figure ci-dessous représente l'architecture de la solution de supervision :

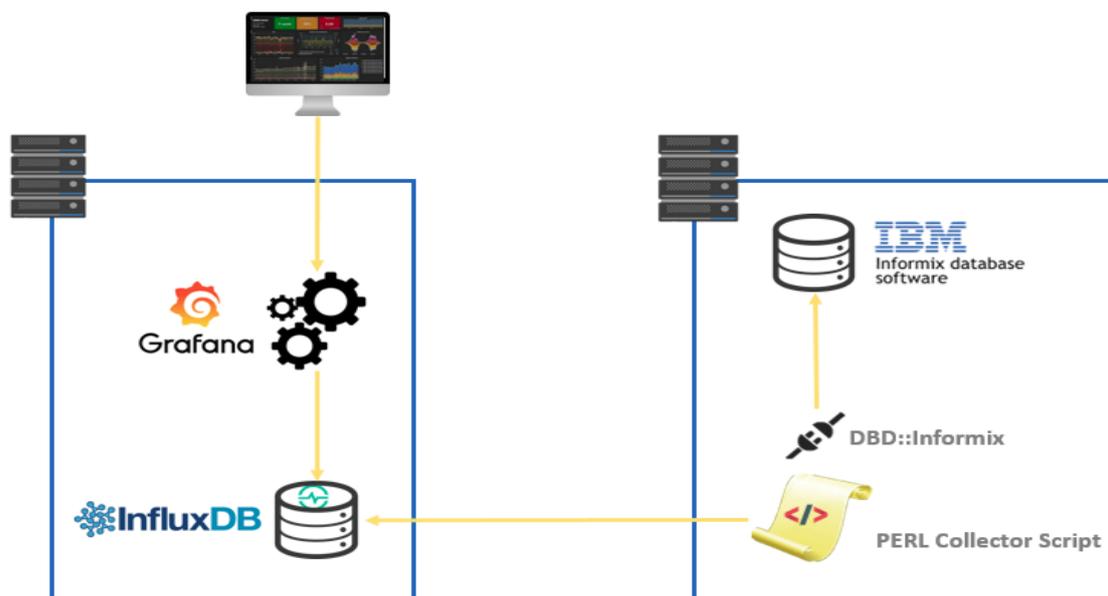


Figure 8: Architecture de la solution de supervision

5. Présentation des technologies utilisées

Une solution de monitoring performante et efficace se doit d'utiliser des technologies à la fois modernes et performantes, c'est pour cela que nous avons opté pour **InfluxDB** une base de données optimisée pour le stockage d'événements et **Grafana** pour la visualisation.

5.1. InfluxDB

5.1.1 Définition

InfluxDB est une base de données de séries temporelles conçue pour gérer des charges d'écriture et de requête élevées. C'est un composant intégral du stack (pile) **TICK**, Telegraf InfluxDB Chronograf Kapacitor. [26]

Time Series Database ou série temporelle est une série de points qui représente une série de mesures identifiées par l'instant auquel elles ont été prises. [27]

5.1.2 Principales caractéristiques

Ci-dessous quelques fonctionnalités actuellement prises en charge par InfluxDB qui en font un excellent choix pour travailler avec des données de séries temporelles : [26]

- Datastore haute performance personnalisé écrit spécifiquement pour les données de séries temporelles.
- Écrit entièrement en Go, un langage de programmation open source, permettant de créer facilement des logiciels simples, fiables avec une efficacité optimale.
- API HTTP d'écriture et de requête simples et performantes.
- Les plugins prennent en charge d'autres protocoles d'intégration de données tels que Graphite, collectd et OpenTSDB.
- Langage de requête expressif de type SQL conçu pour interroger facilement des données agrégées.
- La politique de rétention (RP) qui est la partie de la structure de données d'InfluxDB qui décrit pendant combien de temps InfluxDB conserve les données.

5.1.3 Concepts clés

- Une base de données comporte généralement plusieurs **mesures**. **Une mesure** est à peu près équivalente à une table relationnelle, mais elle n'impose aucune structure aux données stockées.
- Tout comme les tables relationnelles contiennent des enregistrements, **les mesures** contiennent des **points**.
- Chaque **point** a des **champs** obligatoires et un horodatage.
- Un point peut également avoir un ou plusieurs tags facultatifs qui permettent de qualifier les valeurs mesurées.
- **Un horodatage ou TIMESTAMP** spécifie quand son point a été créé. C'est un indice primaire d'une mesure.
- **Un champ** est une paire clé/valeur, qui est similaire à une colonne avec sa valeur dans une table relationnelle.

5.2. Grafana

5.2.1 Définition

Grafana est une suite analytique et de visualisation de métriques open source. Le plus souvent utilisé pour visualiser des données de séries chronologiques pour l'analyse des infrastructures et des applications, mais beaucoup l'utilisent dans d'autres domaines, notamment **les capteurs industriels, la domotique, la météo et le contrôle des processus**. Grafana expose dans des dashboards les métriques brutes ou agrégées provenant de InfluxDB ou d'autres sources de données et permet de définir de manière simple des seuils d'alertes et les actions associées. [27]

Notre choix s'est porté sur Grafana en termes de visualisation, de création de tableaux de bord et de personnalisation pour les raisons ci-dessous :

- Un outil open source.
- Un outil riche en fonctionnalités, facile à utiliser et très flexible.
- Excellent pour créer des graphiques en temps réel de données numériques et temporelles.
- Compatible avec plusieurs types de base de données, dont InfluxDB.

5.2.2 Architecture

La figure ci-dessous présente l'architecture de Grafana : [28]

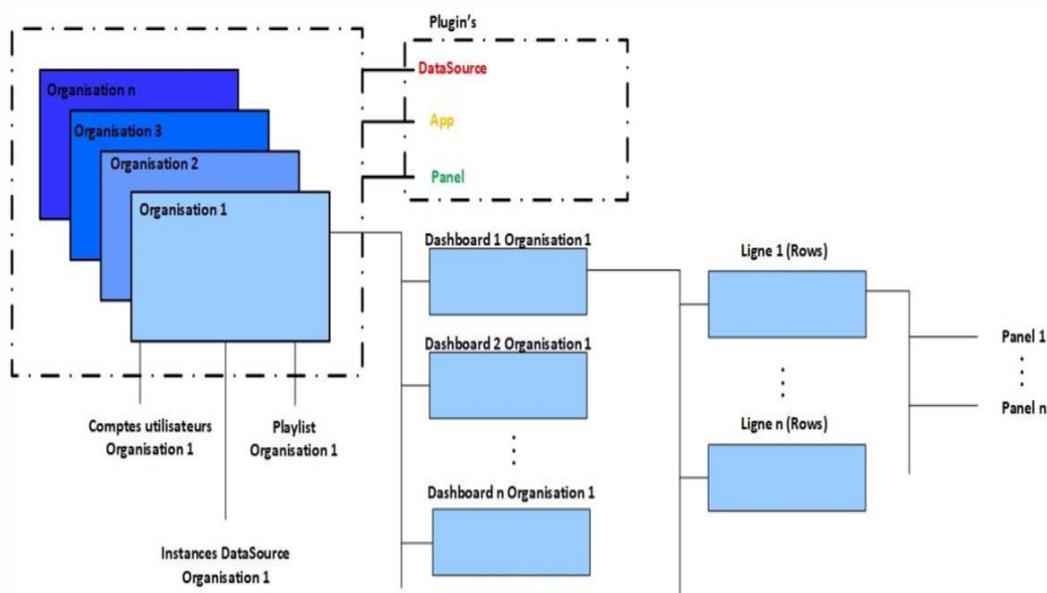


Figure 9: Architecture de Grafana

A. Les Organisations

Ce sont des entités dédiées et cloisonnées sur lesquelles nous pouvons définir :

- Des comptes utilisateurs et leurs droits.
- La connexion aux datasources pour un panel donné.
- La mise en forme des données collectées grâce aux différents panels à notre disposition.

Chaque organisation peut avoir une ou plusieurs sources de données.

Tous les tableaux de bord appartiennent à une organisation particulière.

B. Plugins

Le cœur s'appuie sur trois types de plugins :

a) Les dataSources

Grafana prend en charge de nombreux backends de stockage différent pour les données de séries temporelles (Data Source). Chaque source de données à un éditeur de requête spécifique qui est personnalisé pour les fonctions et les fonctionnalités exposées par la source de données particulière.

Les sources de données suivantes sont officiellement supportées :

- ✓ Graphite
- ✓ Elasticsearch
- ✓ CloudWatch
- ✓ InfluxDB
- ✓ OpenTSDB
- ✓ Prometheus
- ✓ MySQL
- ✓ Microsoft SQL Server (MSSQL)

Le langage de requête et les capacités de chaque source de données sont évidemment très différents. On peut combiner des données provenant de plusieurs sources de données sur un seul tableau de bord, mais chaque panel est lié à une source de données spécifique appartenant à une organisation particulière.

L'interface graphique de Grafana nous permet facilement de choisir entre les différents datasources installés, ce qu'on peut voir dans la figure ci-dessous :

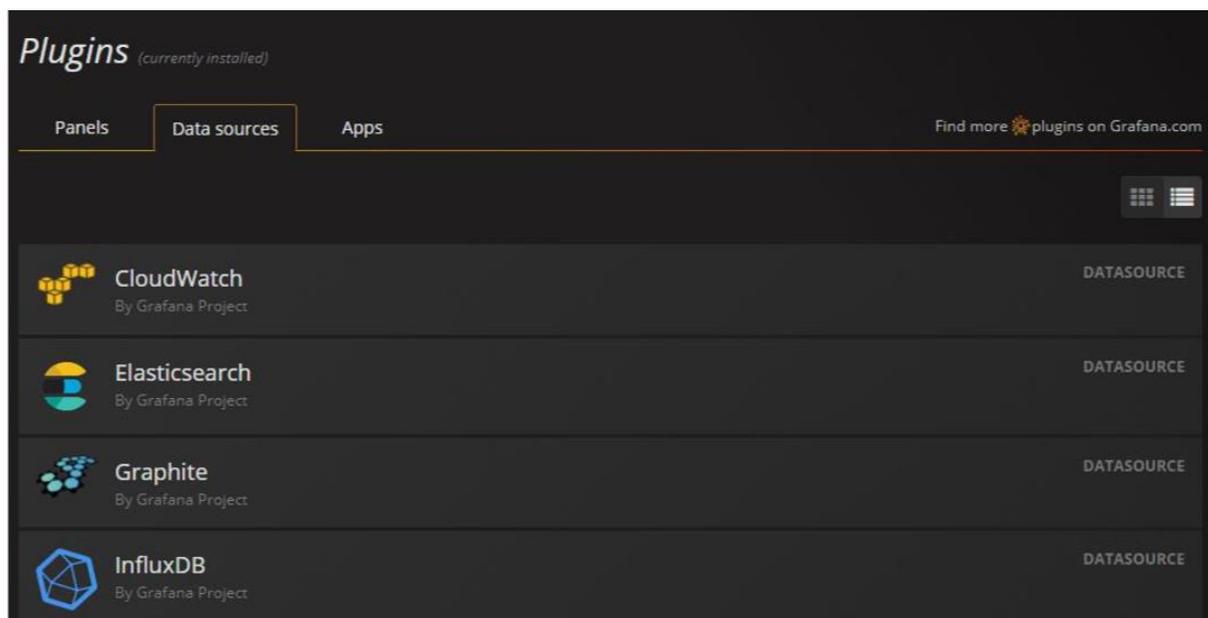


Figure 10: Sources de données dans Grafana

b) Les panels

Un panel est un plugin que fournit Grafana qui est un élément de base de la visualisation des différentes métriques collectées.

Chaque panel fournit un éditeur de requête (en fonction de la source de données sélectionnée dans le panel) qui permet d'extraire la visualisation parfaite à afficher sur le panel en utilisant l'éditeur de requête.

Il existe une grande variété d'options de mise en forme et de style que chaque panel expose pour permettre de créer une image parfaite.

On distingue : les graphiques, les tables, les singlstats, text, etc.

Dans ce qui suit nous expliquons en détails chaque étape de cette démarche présentée.

Etape 1 : La collecte de métriques

Construire un moteur efficace pour stocker et gérer des séries chronologiques est l'un des défis d'un système de surveillance. Comment collecter et stocker des informations à partir d'une source est également important.

Il existe plusieurs sources différentes pour collecter des données : applications, matériels, machines virtuelles, bases de données, etc.

Les techniques les plus utilisées sont les agents collecteurs tels que Collectd et Telegraf.

Dans notre cas nous allons pouvoir récupérer nos métriques à partir de notre base de données Informix avec un script écrit en langage Perl.

Etape 2 : Le stockage

Une fois les métriques collectées à partir de notre base de données Informix nous aurons besoin de les stocker dans une base de données de séries chronologique InfluxDB.

Etape 3 : la visualisation

Après avoir stocké les métriques dans notre base de données InfluxDB, nous allons utiliser l'outil Grafana qui va nous permettre de mettre en forme ces dernières et de générer des graphiques pour visualiser l'état de notre base de données monitorée.

Etape 4 : la remontée d'information

Une fois les visualisations générées nous procédons à l'étape de mise en place d'alertes.

Nous commençons par créer une alerte après avoir définis les indicateurs les plus intéressants que nous souhaitons surveiller.

Ensuite nous devons configurer la ou les conditions de déclenchement de celle-ci.

Enfin pour être prévenu en cas de dépassement des seuils définis, Grafana propose de nombreux types de notifications, dans notre cas nous notifions par mail.

7. Conclusion

A travers ce chapitre nous avons présenté en détail les technologies utilisées pour notre solution de supervision et avons expliqué les étapes de sa réalisation.

Le chapitre suivant entamera l'aspect technique de notre projet ainsi que sa mise en place et son fonctionnement.

Chapitre IV : Réalisation

1. Introduction

Dans ce chapitre, nous nous intéressons à la description de la phase de réalisation de notre application. Nous commençons par la spécification des différents environnements de développement, matériels et logiciels.

Nous présentons par la suite les différentes étapes d'installation et configuration des outils utilisés pour la mise en place de notre plateforme de supervision.

2. Environnement de travail

2.1. Environnement matériel

Durant le stage en entreprise, nous avons utilisé une machine comme environnement de test dont la configuration est la suivante :

- PC portable HP Pavilion
- RAM : 4 Go
- Processeur : Intel Core I3

Une fois notre solution développée, testée et recettée. L'équipe de production envisage de la mettre en production sur un serveur dédié au monitoring, ayant un flux direct avec le serveur contenant la base de données à monitorer.

2.2. Environnement logiciel

2.2.1 Système d'exploitation

Le système d'exploitation utilisé sur l'environnement de test, et plus tard sur celui de production est CentOS 7.

CentOS 7(Community enterprise Operating System 7) est une distribution GNU/Linux, complètement Open Source.

3. Architecture de la plateforme

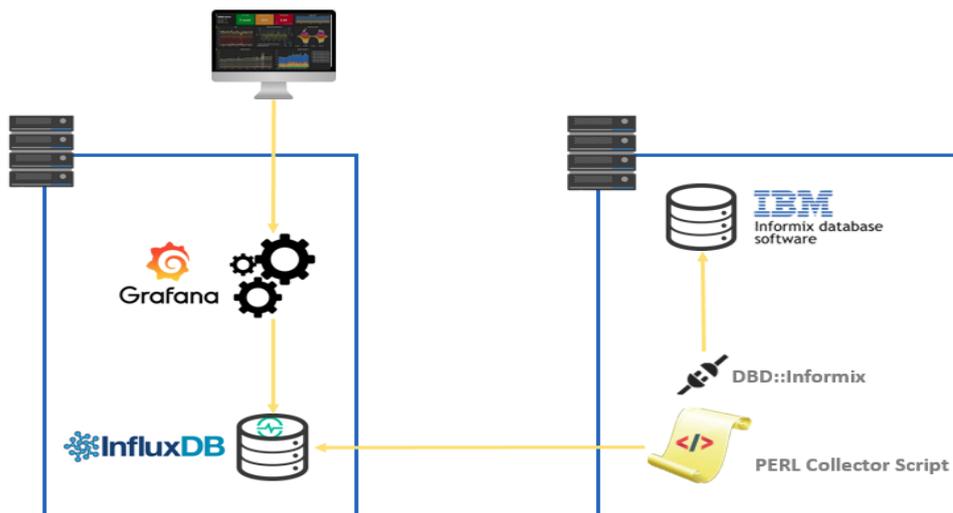


Figure 12: Architecture de la solution

4. Fonctionnement

Cette plateforme supervise en temps réel la base données Informix, commençant par collecter les métriques de cette dernière, les stockant dans une base de données chronologique InfluxDB en utilisant un script codé en Perl.

Le script se connecte à la base de données Informix à travers le driver DBD::Informix, puis alimente la base de données InfluxDB.

La représentation graphique des métriques que nous choisissons se fera sous l'outil de visualisation Grafana, celui-ci utilise la base de données InfluxDB comme source de données pour créer les différents graphiques de notre tableau de bord et permet d'alerter les personnes concernées en cas de dysfonctionnement.

5. Installation et configuration

Nous allons à présent, détailler l'installation des outils nécessaires sur l'environnement de test sous le système d'exploitation linux CentOS 7.

5.1. Installation de InfluxDB

Nous allons commencer par l'installation de la base de données InfluxDB qui représente le cœur de notre système, celle-ci contiendra l'ensemble de nos métriques.

L'installation s'effectue comme suit :

- Ajout du répertoire à la configuration de yum

```
[root@shobohost informix]# cat <<EOF | sudo tee /etc/yum.repos.d/influxdb.repo
> [influxdb]
> name = InfluxDB Repository - RHEL {$releasever}
> baseurl = https://repos.influxdata.com/rhel/{$releasever}/{$basearch}/stable
> enabled = 1
> gpgcheck = 1
> gpgkey = https://repos.influxdata.com/influxdb.key
> EOF
```

Une fois le répertoire ajouté à la configuration de yum, l'installation et le démarrage d'InfluxDB peuvent s'effectuer en exécutant les commandes suivantes :

```
yum install influxdb
service influxdb start
```

- Nous commençons par la commande : yum install influxdb

```
[root@shobohost informix]# yum install influxdb
Loaded plugins: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
 * base: mirror.centos.org
 * extras: mirror.centos.org
 * updates: mirror.centos.org
influxdb | 2.5 kB 00:00
influxdb/7/x86_64/primary_db | 29 kB 00:01
Resolving Dependencies
--> Running transaction check
--> Package influxdb.x86_64 0:1.5.2-1 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

=====
Package Arch Version Repository Size
=====
Installing:
influxdb x86_64 1.5.2-1 influxdb 21 M
Transaction Summary
=====
Install 1 Package

Total download size: 21 M
Installed size: 21 M
Is this ok [y/d/N]:
```

Une fois l'installation terminée, on exécute la commande de lancement de notre base de données Influxdb :

```
Installed:
influxdb.x86_64 0:1.5.2-1

Complete!
[root@shobohost informix]# service influxdb start
Redirecting to /bin/systemctl start influxdb.service
[root@shobohost informix]#
```

InfluxDB est désormais installé et exécuté !

5.2. Installation de Grafana

Nous procédons à l'installation manuelle de Grafana sur CentOS 7 en utilisant rpm, nous commençons par télécharger Grafana en version stable :

```
[root@shobohost informix]# wget https://s3-us-west-2.amazonaws.com/grafana-releases/release/grafana-5.1.3-1.x86_64.rpm
--2018-06-14 12:54:02-- https://s3-us-west-2.amazonaws.com/grafana-releases/release/grafana-5.1.3-1.x86_64.rpm
Resolving s3-us-west-2.amazonaws.com (s3-us-west-2.amazonaws.com)... 52.218.160.44
Connecting to s3-us-west-2.amazonaws.com (s3-us-west-2.amazonaws.com)|52.218.160.44|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 52210696 (50M) [application/x-redhat-package-manager]
Saving to: 'grafana-5.1.3-1.x86_64.rpm.2'

100%[=====>] 52,210,696 562KB/s in 97s

2018-06-14 12:55:40 (526 KB/s) - 'grafana-5.1.3-1.x86_64.rpm.2' saved [52210696/52210696]
```

Une fois le téléchargement terminé, il est nécessaire d'installer les dépendances requises pour l'installation de Grafana, pour ensuite procéder à son installation.

```
[root@shobohost informix]# yum install initscripts fontconfig
Loaded plugins: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
 * base: mirror.centos.org
 * extras: mirror.centos.org
 * updates: mirror.centos.org
Package initscripts-9.49.41-1.el7.x86_64 already installed and latest version
Package fontconfig-2.10.95-11.el7.x86_64 already installed and latest version
Nothing to do
[root@shobohost informix]# rpm -Uvh grafana-5.1.3-1.x86_64.rpm
warning: grafana-5.1.3-1.x86_64.rpm: Header V4 RSA/SHA1 Signature, key ID 24098c
b6: NOKEY
Preparing... ##### [100%]
Updating / installing...
 1:grafana-5.1.3-1 ##### [100%]
### NOT starting on installation, please execute the following statements to con
figure grafana to start automatically using systemd
sudo /bin/systemctl daemon-reload
sudo /bin/systemctl enable grafana-server.service
### You can start grafana-server by executing
sudo /bin/systemctl start grafana-server.service
POSTTRANS: Running script
```

Une fois installé, il suffit d'entrer les commandes suivantes pour démarrer le service :

```
systemctl daemon-reload
systemctl enable grafana-server
systemctl start grafana-server
```

```
[root@shobohost informix]# systemctl daemon-reload
[root@shobohost informix]# systemctl enable grafana-server
Created symlink from /etc/systemd/system/multi-user.target.wants/grafana-server.
service to /usr/lib/systemd/system/grafana-server.service.
[root@shobohost informix]# systemctl start grafana-server
[root@shobohost informix]# systemctl status grafana-server
● grafana-server.service - Grafana instance
   Loaded: loaded (/usr/lib/systemd/system/grafana-server.service; enabled; vend
or preset: disabled)
   Active: active (running) since Thu 2018-06-14 13:18:41 CET; 27s ago
     Docs: http://docs.grafana.org
   Main PID: 30606 (grafana-server)
    Tasks: 13
   CGroup: /system.slice/grafana-server.service
           └─30606 /usr/sbin/grafana-server --config=/etc/grafana/grafana.ini...

Jun 14 13:18:41 shobohost grafana-server[30606]: t=2018-06-14T13:18:41+0100 ..."
Jun 14 13:18:41 shobohost grafana-server[30606]: t=2018-06-14T13:18:41+0100 ..."
Jun 14 13:18:41 shobohost grafana-server[30606]: t=2018-06-14T13:18:41+0100 ..."
Jun 14 13:18:41 shobohost grafana-server[30606]: t=2018-06-14T13:18:41+0100 ...s
Jun 14 13:18:41 shobohost grafana-server[30606]: t=2018-06-14T13:18:41+0100 ...s
Jun 14 13:18:41 shobohost grafana-server[30606]: t=2018-06-14T13:18:41+0100 ...e
Jun 14 13:18:41 shobohost grafana-server[30606]: t=2018-06-14T13:18:41+0100 ...p
Jun 14 13:18:41 shobohost systemd[1]: Started Grafana instance.
Jun 14 13:18:41 shobohost grafana-server[30606]: t=2018-06-14T13:18:41+0100 ..."
Jun 14 13:18:41 shobohost grafana-server[30606]: t=2018-06-14T13:18:41+0100 ...=
Hint: Some lines were ellipsized, use -l to show in full.
[root@shobohost informix]#
```

Le serveur est donc accessible sur le port 3000 de notre machine. Ci-dessous la page de connexion Grafana depuis un navigateur web.

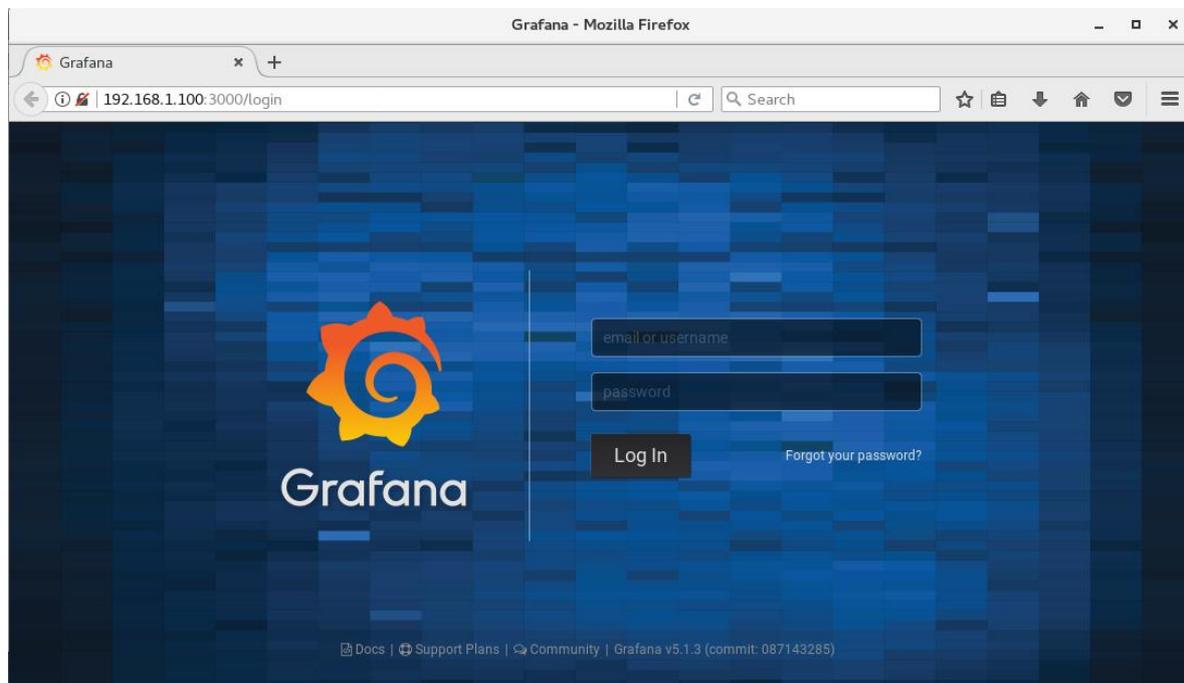


Figure 13: Connexion à Grafana

5.3. Mise à jour de Perl

Perl est une famille de langages de programmation de scripts dont la syntaxe est similaire à celle du langage C, y compris Perl 5 et Perl 6. Perl est un langage open source, interprété et à usage général. [29]

Perl est installé par défaut sur CentOS 7, il suffit de le mettre à jour.

5.4. Installation de DBD::Informix

DBD::informix est le driver permettant au langage Perl de se connecter à une base de données Informix, et ainsi pouvoir lancer des requêtes sur celle-ci depuis un programme (script) Perl.

Pour pouvoir installer le driver perl de la base de données Informix (DBD::informix), nous aurons besoins de :

A. CPAN

Acronyme de Comprehensive Perl Archive Network, est une collection d'archives Internet constituant une ressource au langage de programmation Perl. CPAN comprend : [30]

- Une liste des modules Perl et leur code source.
- Documentation pour Perl et les modules Perl.
- Distributions de Perl.
- Informations sur Perl et les auteurs de modules Perl.
- Conseils pour la programmation en Perl.

Dans notre cas nous utilisons CPAN pour télécharger, compiler, tester et installer Perl DBI et DBD::Informix, des modules de Perl.

B. Perl DBI

DBI est un module d'accès à une base de données pour le langage de programmation Perl. Il définit un ensemble de méthodes, de variables et de conventions qui fournissent une interface de base de données cohérente, indépendante de la base de données réelle utilisée. [31]

C. Définition des variables d'environnement

Pour que le driver DBD:informix s'installe correctement, et pour qu'il puisse effectuer les premiers tests de connexion, il est nécessaire de déclarer certaines variables d'environnement en relation avec notre base de données informix :

```
export DBI_DBNAME=sysmaster
export DBD_INFORMIX_DATABASE=sysmaster
export DBD_INFORMIX_USERNAME=informix
export DBD_INFORMIX_PASSWORD=informix
export LD_LIBRARY_PATH=$INFORMIXDIR/lib:$INFORMIXDIR/lib/esql:$INFORMIXDIR/lib/tools;
```

D. Initialisation des variables d'environnement

En plus des variables d'environnement déclarées dans le point précédent, il convient de charger et d'initialiser les variables d'environnement stockées dans un fichier de configuration de la base de données Informix :

```
[informix@shobohost ~]$ . /home/informix/ifx_dev.env
```

Une fois les prérequis établis, nous allons procéder à l'installation de DBD::Informix comme le montre la figure suivante :

```
[informix@shobohost ~]$ cpan
Terminal does not support AddHistory.

cpan shell -- CPAN exploration and modules installation (v1.9800)
Enter 'h' for help.

cpan[1]>
Reading '/home/informix/.cpan/Metadata'
  Database was generated on Tue, 26 Jun 2018 10:41:02 GMT
Running install for module 'DBD::Informix'
Running make for J/J0/JOHNL/DBD-Informix-2015.1101.tar.gz
Checksum for /home/informix/.cpan/sources/authors/id/J/J0/JOHNL/DBD-Informix-2015.1101.tar.gz ok
Scanning cache /home/informix/.cpan/build for sizes
.....DONE

  CPAN.pm: Building J/J0/JOHNL/DBD-Informix-2015.1101.tar.gz

*** ExtUtils::AutoInstall version 0.63
*** Checking for dependencies...
[Core Features]
- DBI      ...loaded. (1.641 >= 1.607)
[High Resolution Timing]
- Time::HiRes ...loaded. (1.9725)
```

6. Mise en place de la plateforme de supervision

6.1. Création de la base de données Influxdb

Dans cette étape nous allons créer la base de données chronologique InfluxDB, où seront stockées les métriques collectées depuis la base de données monitorée Informix :

```
[informix@shobohost ~]$ influx
Connected to http://localhost:8086 version 1.5.2
InfluxDB shell version: 1.5.2
> create database informix
> use informix
Using database informix
>
```

6.2. La collecte de métriques

Comme déjà expliqué auparavant, nos tableaux de bord vont avoir comme source, la base de données chronologique InfluxDB, celle-ci sera alimenté par un script Perl, qui procèdera comme suit :

- Le script utilise DBD:informix pour établir une connexion à la base de données Informix
- Collecter toutes les métriques depuis Informix en utilisant des requêtes prédéfinies.
- Mettre en forme les résultats des requêtes pour correspondre aux caractéristiques de la base de données influxDB.
- Préparation et exécution des requêtes POST, permettant d'écrire sur la base de données InfluxDB.

Toutes ces tâches se répètent continuellement toutes les 10 secondes.

6.3. Connexion Grafana influxdb

Nous commençons alors par configurer la connexion à la source de données InfluxDB.

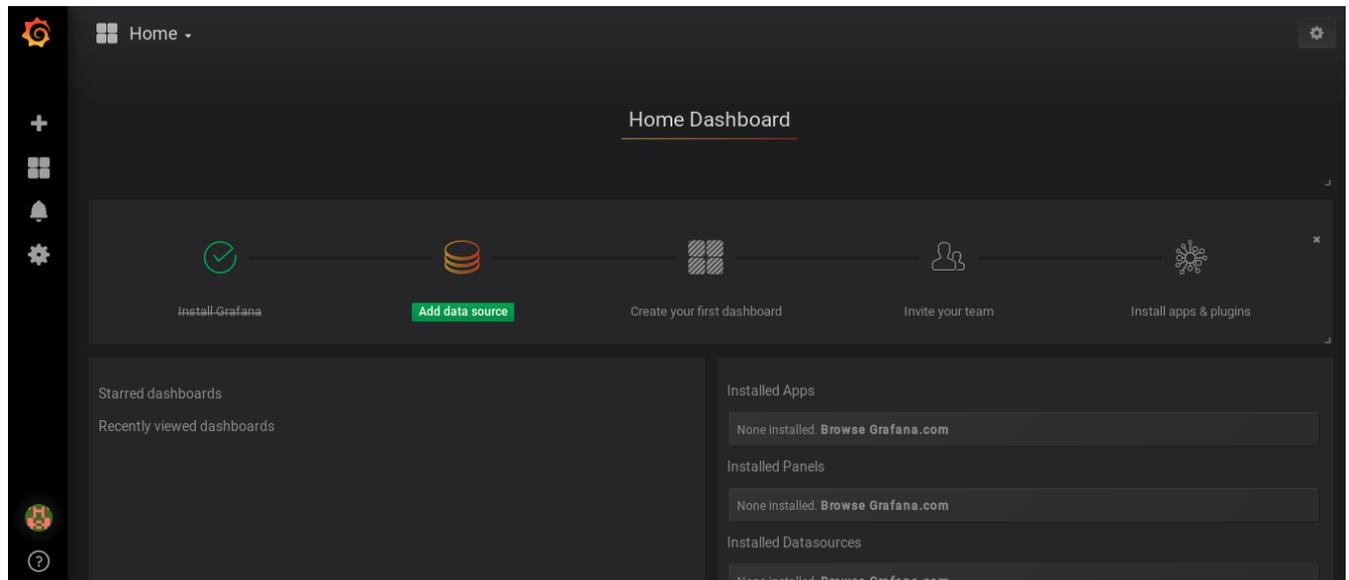


Figure 14 : Ajout d'une source de données

On suit la configuration suivante :

- Comme la base de données à monitorer est Informix, on choisit de nommer cette connexion par : **Informix**.
- Le type de source de données à laquelle on va se connecter est **InfluxDB**.
- Selon la configuration de la base de données et du port d'écoute, on configure l'url de connexion à celle-ci avec le modèle
 - **http://<ip_bdd>:<port>**

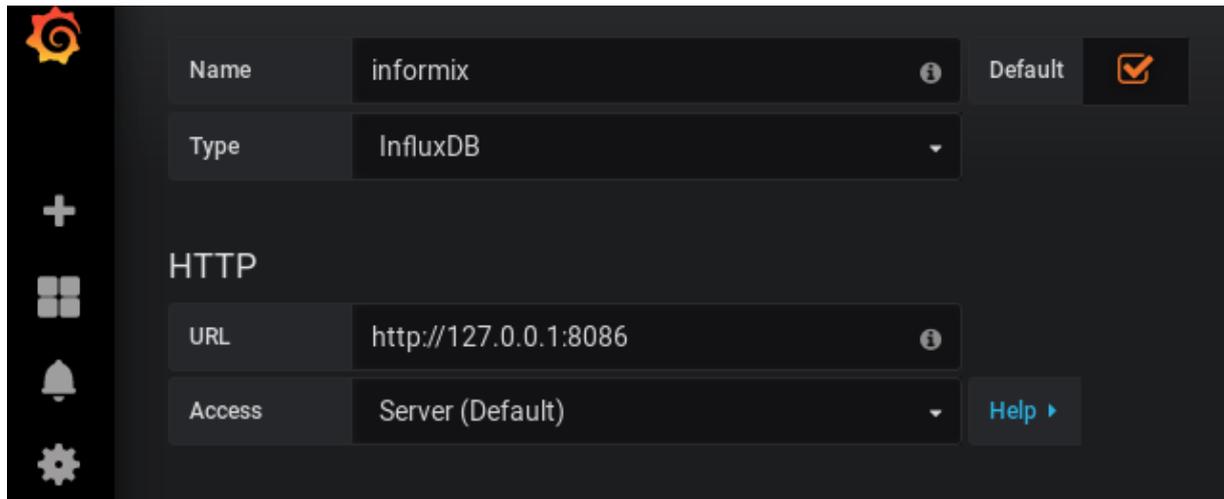


Figure 15: Configuration de la source de données

- Dans le pavé **InfluxDB Details** on mettra le nom de la base de données créée précédemment lors de la configuration d’InfluxDB (dans notre exemple la base de données qui reçoit les mesures sur InfluxDB nommée Informix).
- Une fois ces configurations achevées, on sauvegarde et teste la connexion.

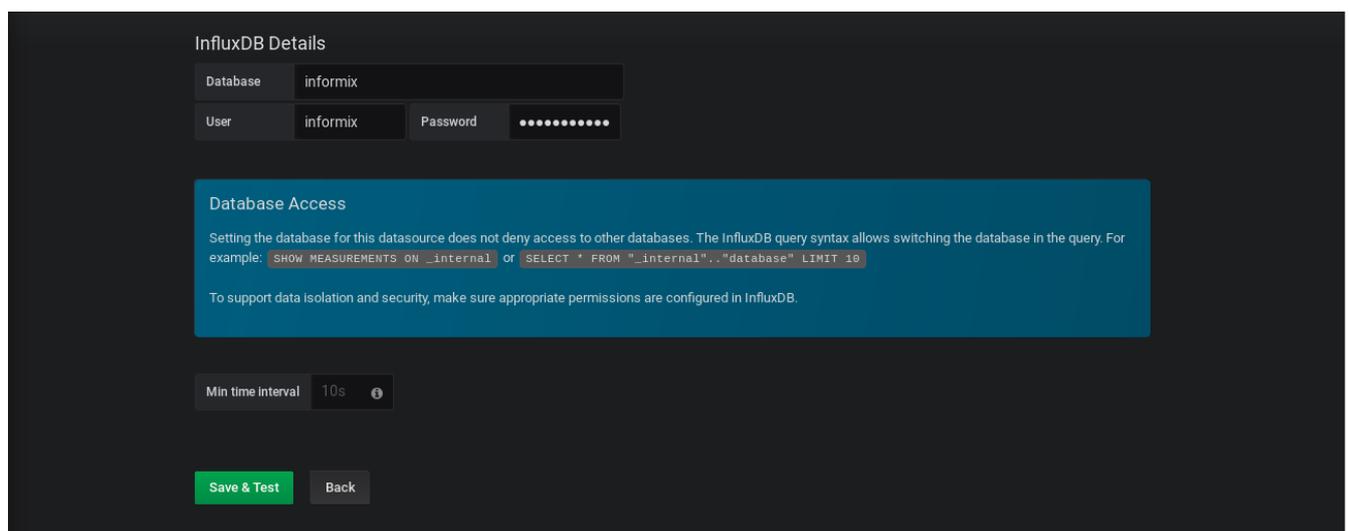


Figure 16: Test de la connexion à la source de données

6.4. Création de visualisations

Une fois la connexion entre Grafana et la source de données InfluxDB établie, nous pouvons à présent créer des visualisations qui constitueront notre tableau de bord.

Nous pouvons distinguer différents indicateurs de performance des bases de données Informix, ci-dessous quelques exemples réalisés :

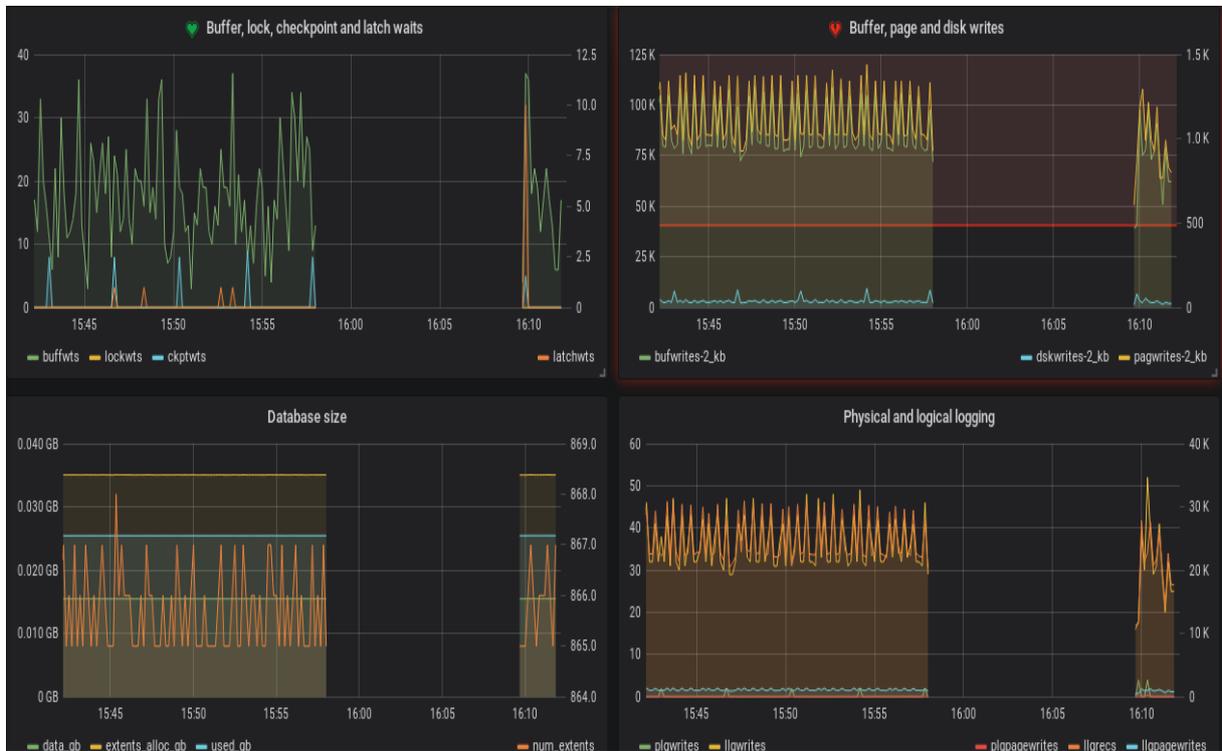


Figure 17: Exemple de dashboard

Exemple 1 : La visualisation suivante permet d'avoir des informations sur la taille de la base de données.

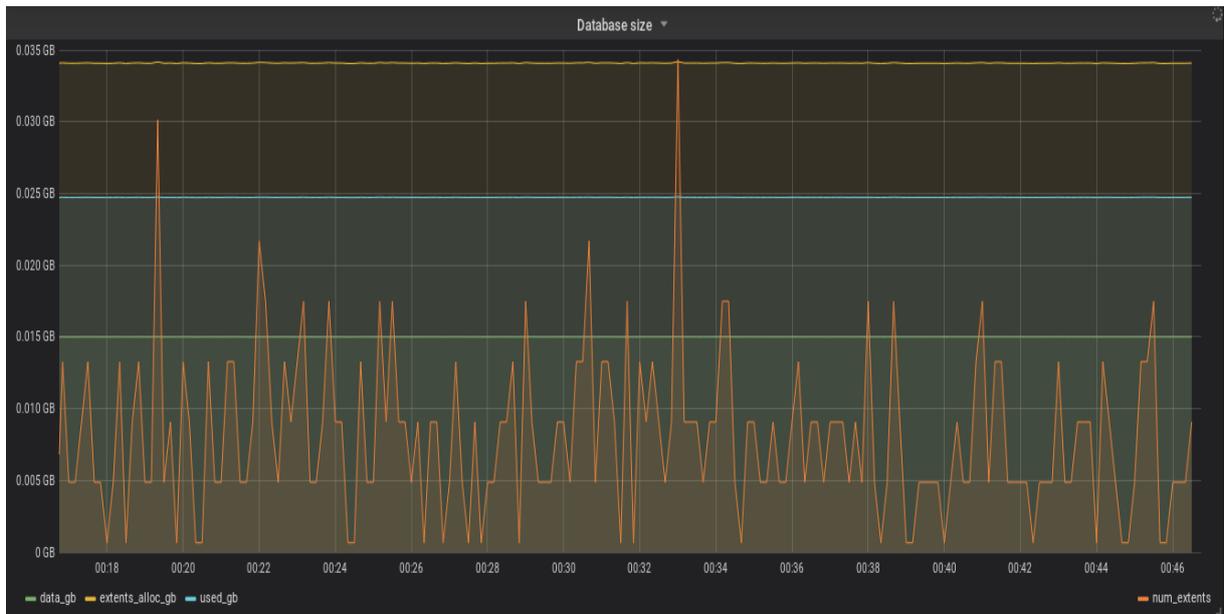


Figure 18: Database size

Exemple 2

La visualisation suivante montre trois graphes qui représentent des métriques récupérées depuis la source de données influxDB qui sont **bufwrits**, **diskwrits** et **pagewrits**, cela permet de surveiller le nombre d'écriture dans le buffer, le nombre de pages écrites sur le disque à partir de celui-ci ainsi que le nombre d'opérations d'écriture sur disque effectuées pour réécrire les pages modifiées de nouveau sur le disque.



Figure 19: Buffer,page and disk writes

Exemple 3

Le graphe ci-dessous représente les métriques **latch_wait**, **lock_wait**, **bufwaits** et **checkpoint_wait** qui permettent de surveiller respectivement le nombre de fois qu'un thread a dû attendre un verrou, le nombre de fois que les sessions doivent attendre la libération de celui-ci, Le nombre de fois qu'un thread a dû attendre un verrou sur un buffer ainsi que le nombre de fois que les threads ont dû attendre les points de contrôle de la section critique.

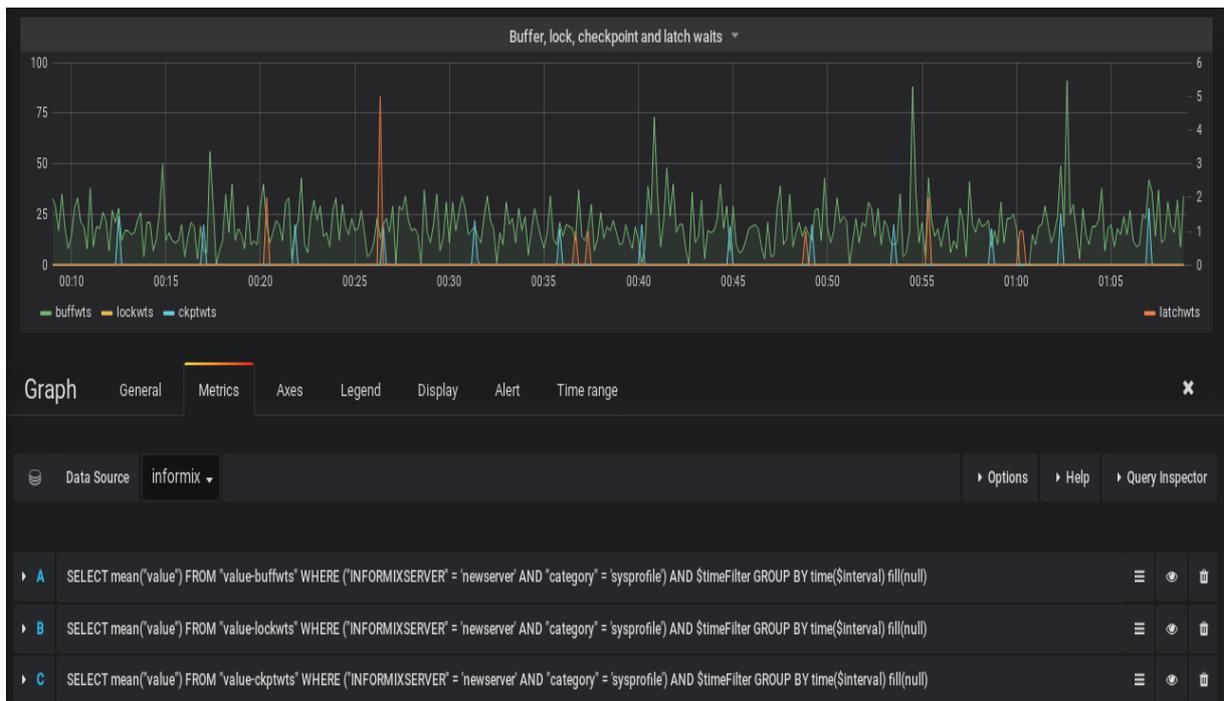


Figure 20: Buffer, lock, checkpoint and latch waits

Exemple 4

Dans cet exemple nous surveillons le journal physique pour déterminer le pourcentage de celui-ci utilisé avant qu'un point de contrôle ne se produise. Ces informations sont utiles pour configurer la taille du fichier journal physique.

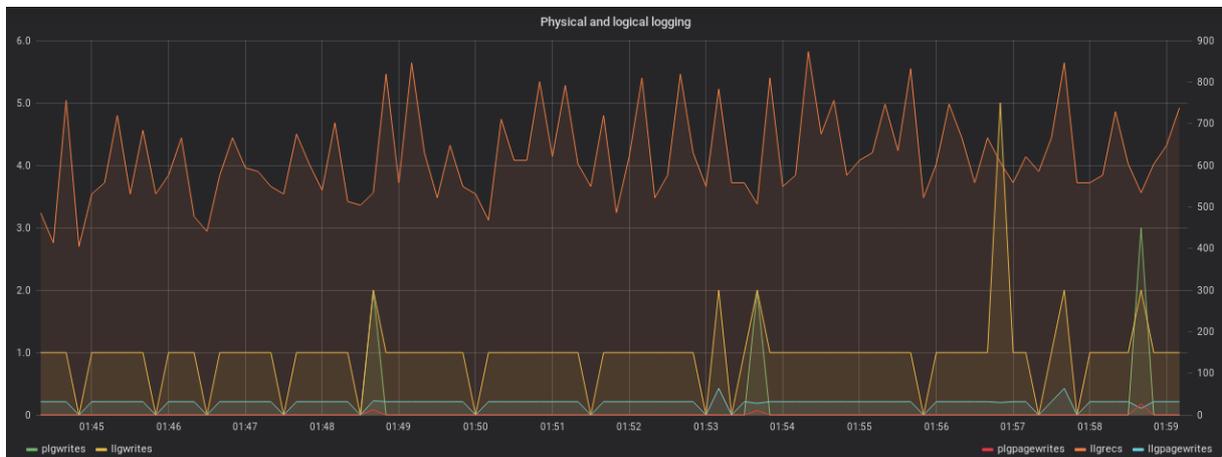


Figure 21: Physical and logical logging

6.5. Mise en place d'alertes

Un système de supervision, permet d'avoir une vision globale de l'état du système supervisé, ceci permet aux personnes chargées de l'exploitation de ce dernier de faire un suivi en temps réel et de pouvoir intervenir en cas d'anomalies constatées sur l'un de ces graphes. Cependant le système doit pouvoir en plus de la mise en disposition de ces visualisations, d'alerter l'exploitant en cas d'anomalies, en surveillant les graphes sachant un seuil critique que l'exploitant aura configuré sur le système de supervision. Pour cela nous utilisons les alertes mis à disposition sur notre solution Grafana.

Pour mettre en places ces alertes nous devons commencer par configurer le canal d'envoi des alertes, dans notre cas le canal Mail sera utilisé.

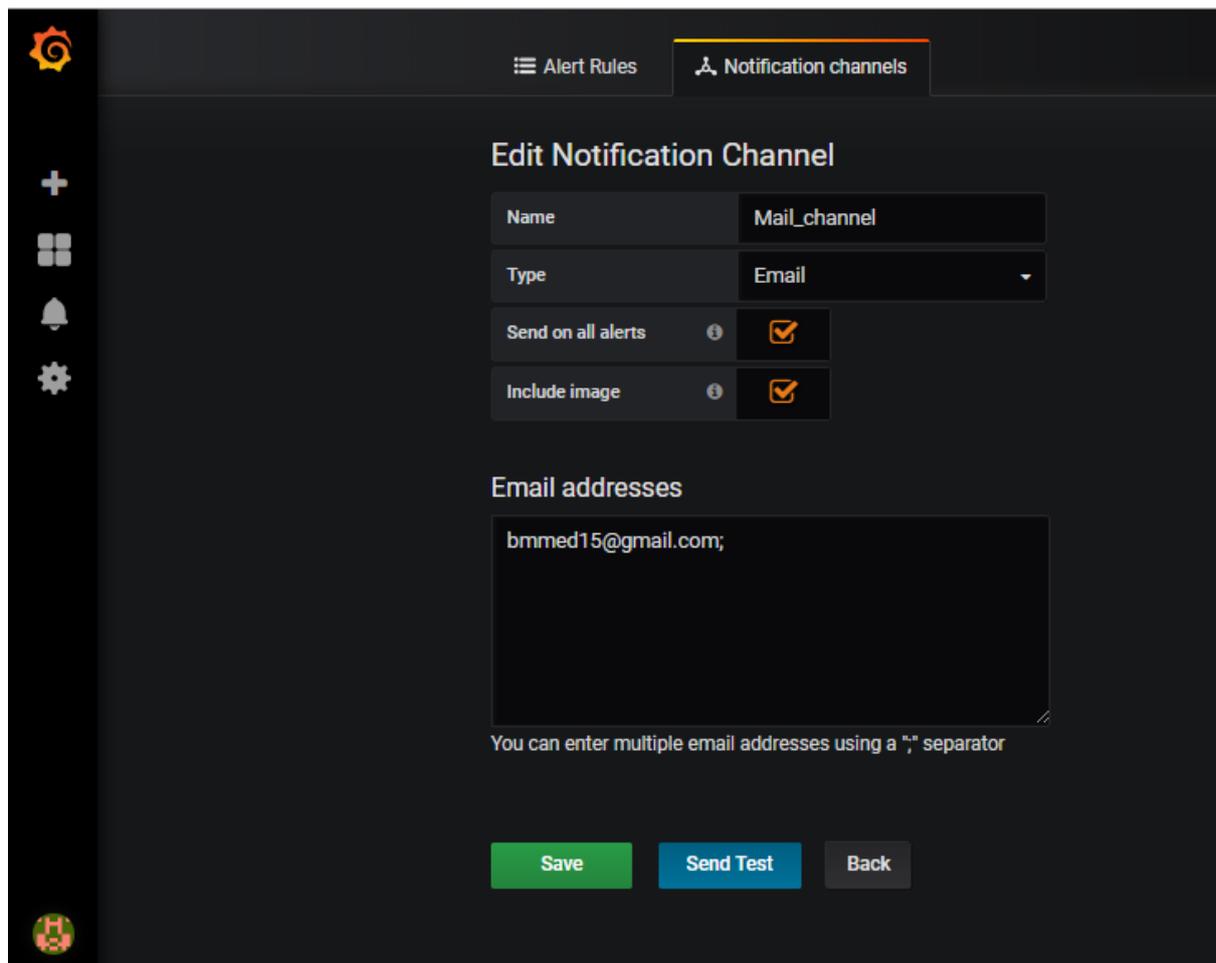


Figure 22 : Configuration du canal de notification

Une configuration des données de connexion au server SMTP est requise, ces données sont à renseigner dans [grafana-5.1.0\conf\defaults.ini](#) sous Windows ou dans [etc/grafana/conf.ini](#) sous linux comme suit :

```
##### SMTP / Emailing #####
[smtplib]
enabled = true
host = smtp.gmail.com:587
user = noreplay.localhost@gmail.com
password = <pass_word>
;cert_file =
;key_file =
skip_verify = true
from_address = noreplay.localhost@gmail.com
```

Figure 23: Configuration du SMTP

Dans notre exemple nous avons utilisé un compte Gmail comme adresse d'envoi de nos alertes, passant par le SMTP Gmail smtp.gmail.com:587.

Une fois notre canal d'envoi d'alerte paramétré, on configure le déclencheur d'alerte sur le graphe que nous souhaitons surveiller.

Exemple

L'exemple ci-dessous montre une configuration d'alerte définie sur le graphe **bufwaits** qui illustre le nombre de fois qu'un thread a dû attendre un verrou sur un buffer, quand le seuil défini est dépassé, l'alerte se déclenche et les personnes concernées seront notifiées par la réception d'un mail.

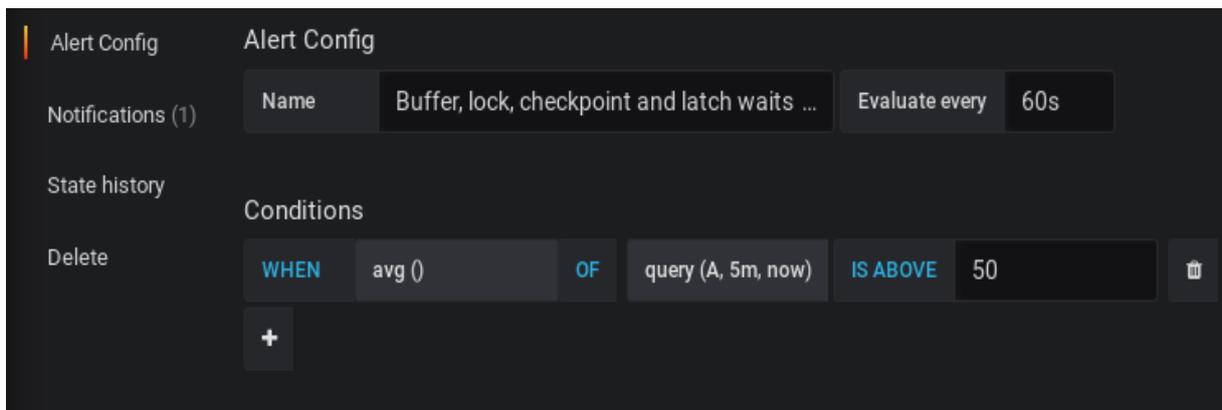


Figure 24: Exemple d'alerte

7. Conclusion

Dans ce chapitre nous avons mis l'intérêt sur l'aspect pratique de notre projet tout en expliquant son fonctionnement, nous avons présenté l'environnement de travail et les différentes étapes d'installation et configuration des outils utilisés. Enfin nous sommes parvenus à créer des dashboards visuels permettant de refléter l'état de notre base de données.

Annexe

1. Installation de la base de données

1.1 Préparation de l'environnement

Nous commençons par la création d'un group utilisateur Informix comme suit :

```
groupadd -g n informix
```

n un nombre entier > 100

Après avoir créé ce groupe on crée l'utilisateur informix comme suivant :

```
useradd -u n -g informix informix
```

1.2 Installation du serveur de base de données informix IDS

Nous allons présenter dans ce qui suit le détail de l'installation d'IDS (Informix Dynamic server) obtenu sur le site officiel d'IBM. En tant qu'utilisateur informix on exécute les commandes suivantes :

Créer un dossier **où** nous déposons nos fichiers d'installation

```
[informix@shobohost ~]$ cd /home/informix  
[informix@shobohost ~]$ mkdir informix_install
```

```
[informix@shobohost ~]$ cp iif.11.70.FC8DE.linux-x86_64.tar informix_install/
```

Annexe

Une fois tout est mis en place, on procède comme le montre la capture suivante :

- Lancement de l'installation

```
[informix@shobohost ~]$ cd informix_install/
[informix@shobohost informix_install]$ ls -l
total 1013140
-rw-r--r--. 1 informix informix 10201 Jun 10 2017 bundle.properties
drwxr-xr-x. 3 informix informix 17 Jun 10 2017 CSDK
drwxr-xr-x. 11 informix informix 4096 Jun 10 2017 IBM_Data_Server_Driver_P
ackage
drwxr-xr-x. 3 informix informix 17 Jun 10 2017 ICONNECT
-rwxr-xr-x. 1 informix informix 480078869 Jun 10 2017 ids_install
-rw-r--r--. 1 informix informix 557342720 Apr 30 20:56 iif.12.10.FC9TL.linux-x8
6_64.tar
drwxr-xr-x. 5 informix informix 50 Jun 10 2017 JDBC
-r--r--r--. 1 informix informix 13008 Jun 10 2017 README.html
drwxr-xr-x. 3 informix informix 33 Jun 10 2017 SERVER
[informix@shobohost informix_install]$ su
Password:
[root@shobohost informix_install]# ./ids_install
Preparing to install...
Extracting the JRE from the installer archive...
Unpacking the JRE...
Extracting the installation resources from the installer archive...
Configuring the installer for this system's environment...
Launching installer...
```

- Choix du répertoire d'installation

```
=====  
Installation Location  
-----  
  
Choose location for software installation  
  
Default Install Folder: /opt/informix12.10.FC9T  
ENTER AN ABSOLUTE PATH, OR PRESS <ENTER> TO ACCEPT THE DEFAULT  
:  
  
=====  
Installation or Distribution  
-----  
  
Select the installation type.  
  
Typical: Install the database server with all features and a database server  
that  
is configured with default values. Includes:  
** Client Software Development Kit (CSDK)
```

➤ Choix du type d'installation

```
Custom: Install the database server with specific features and software that
you need.
Optionally install a configured database server instance.
Minimum disk space required: 75 MB (without a server instance)

->1- Typical installation
   2- Custom installation
   3- Extract the product files (-DLEGACY option)

ENTER THE NUMBER FOR YOUR CHOICE, OR PRESS <ENTER> TO ACCEPT THE DEFAULT:: 2
```

➤ Choix des fonctionnalités à installer

```
"Choose Install Features" Help
-----

On the next screen, type the number of a feature to change whether it is
selected or deselected. You can provide a list of numbers, separated by commas,
with no leading or trailing spaces.

For example, assume that the screen contains this list:

1- [X] ABC
2- [ ] XYZ

Please choose the Features to be installed by this installer.: ____

If you want to install both ABC and XYZ, type: 2
If you want to install only XYZ, type: 1,2

PRESS <ENTER> TO CONTINUE:
```

```
Please choose the Features to be installed by this installer.:

=====
Feature review
-----

Would you like to review the features that you have selected, or change which
features you have selected?

->1- Continue
   2- Return to feature tree

ENTER THE NUMBER OF THE DESIRED CHOICE, OR PRESS <ENTER> TO ACCEPT THE
DEFAULT:
```

➤ Configuration des rôles

```
Get Role Separation choice
-----

Enable role separation for auditing procedures.

If you enable role separation, you can assign existing groups of users to
specific roles.

If you do not enable role separation, the database server administrator
performs all administration tasks.

   1- Enable role separation
->2- Do not enable role separation

ENTER THE NUMBER FOR YOUR CHOICE, OR PRESS <ENTER> TO ACCEPT THE DEFAULT:: 2
```

- Choix création de l'instance

```
Server Instance
-----
Type 'back' to go to the previous step or 'quit' to cancel the installation.
Create a database server instance?

    1- Yes - create a server instance
    ->2- No - do not create a server instance

ENTER THE NUMBER FOR YOUR CHOICE, OR PRESS <ENTER> TO ACCEPT THE DEFAULT:: 2
```

Annexe

- Installation de informix IDS

```
Installing...
-----

[=====|=====|=====|=====]
[-----|-----|-----|-----]

=====
Installation Complete
-----

Congratulations! IBM Informix Software Bundle installation is complete.

Product install status:
Informix Dynamic Server: Successful
Informix Client-SDK: Successful

For more information about using Informix products, see the IBM Informix 12.10
Information Center at http://pic.dhe.ibm.com/infocenter/informix/v121/index.jsp.
```

```
Disk Space Information (for Installation Target):
  Required: 676,457,500 Bytes
  Available: 45,006,913,536 Bytes

PRESS <ENTER> TO CONTINUE:

=====

Ready To Install
-----

InstallAnywhere is now ready to install IBM Informix Software Bundle onto your
system at the following location:

  /opt/informix12.10.FC9T

PRESS <ENTER> TO INSTALL:
```

1.3 Configuration post installation

Une fois le serveur installé, on doit effectuer les configurations suivantes :

- Définition des variables d'environnement

```
[root@shobohost informix_install]# cat /etc/profile.d/informix.sh
INFORMIXDIR=/opt/informix12.10.FC9T/
export INFORMIXDIR
PATH=$INFORMIXDIR/bin:$PATH
export PATH
INFORMIXSERVER=newserver
export INFORMIXSERVER
ONCONFIG=onconfig.$INFORMIXSERVER
export ONCONFIG
```

- Création des fichiers pour les chunks

```
[root@shobohost informix_install]# mkdir /informixchunks
[root@shobohost informix_install]# cd /informixchunks/
[root@shobohost informixchunks]# mkdir newserver
[root@shobohost informixchunks]# cd newserver
[root@shobohost newserver]# touch /informixchunks/newserver/rootdbs
[root@shobohost newserver]# touch /informixchunks/newserver/logdbs
[root@shobohost newserver]# touch /informixchunks/newserver/datadbs
[root@shobohost newserver]# touch /informixchunks/newserver/tmpdbs
[root@shobohost newserver]# chown informix:informix /informixchunks/
[root@shobohost newserver]# chown informix:informix /informixchunks/newserver/rootdbs
[root@shobohost newserver]# chown informix:informix /informixchunks/newserver/logdbs
[root@shobohost newserver]# chown informix:informix /informixchunks/newserver/datadbs
[root@shobohost newserver]# chown informix:informix /informixchunks/newserver/tmpdbs
[root@shobohost newserver]# █
```

- Configuration de la connectivité

```
[root@shobohost etc]# nano sqlhosts
[root@shobohost etc]# cat sqlhosts
newserver      onipcshm      shobohost     newserver
newservertcp  onsoctcp      shobohost     8089
```

Annexe

Dans le fichier de configuration établie les configurations suivantes :

```
[root@shobohost etc]# cp onconfig.std $ONCONFIG
[root@shobohost etc]# ls -l onco*
-rw-r--r--. 1 root      root      84540 Jul  2 18:57 onconfig.newserver
-rw-r--r--. 1 informix informix 84540 Jun 10 2017 onconfig.std
[root@shobohost etc]# chown informix informix onconfig.newserver
chown: cannot access 'informix': No such file or directory
[root@shobohost etc]# chown informix:informix onconfig.newserver
[root@shobohost etc]# ls -l onco*
-rw-r--r--. 1 informix informix 84540 Jul  2 18:57 onconfig.newserver
-rw-r--r--. 1 informix informix 84540 Jun 10 2017 onconfig.std
[root@shobohost etc]# nano $ONCONFIG
```

```
ROOTNAME rootdbs
ROOTPATH /informixchunks/newserver/rootdbs
ROOTOFFSET 0
ROOTSIZE 2000000
MIRROR 0
MIRRORPATH $INFORMIXDIR/tmp/demo_on.root_mirror
MIRROROFFSET 0
```

➤ Initialisation du serveur

```
[informix@shobohost etc]$ oninit -i
Warning: INFORMIXSQLHOSTS /opt/informix12.10.FC9T//etc/sqlhosts is not owned by
user informix (uid=2000).
Warning: INFORMIXSQLHOSTS /opt/informix12.10.FC9T//etc/sqlhosts is not owned by
group informix (gid=2000).
Your evaluation license will expire on 2018-09-30 00:00:00

This action will initialize IBM Informix Dynamic Server;
any existing IBM Informix Dynamic Server databases will NOT be accessible -
Do you wish to continue (y/n)? y
[informix@shobohost etc]$ onstat -m
Your evaluation license will expire on 2018-09-30 00:00:00

IBM Informix Dynamic Server Version 12.10.FC9TL -- On-Line -- Up 00:01:20 -- 14
8632 Kbytes
```

Conclusion générale

Au terme de ces six mois, nous avons pu mettre en place une solution de supervision pour les bases de données Informix permettant ainsi d'avoir une vue globale sur le fonctionnement de la base et d'alerter en cas d'anomalie.

N'ayant pas assez de connaissances, nous avons dû nous adapter et faire beaucoup de recherches sur la supervision, les bases de données Informix ainsi que sur le métier d'administrateur de bases de données pour acquérir une base de connaissances pour mener à bien ce projet dans son ensemble.

Ce projet a été pour nous une réelle opportunité pour améliorer nos compétences et acquérir une expérience très enrichissante et le stage que nous avons effectué nous a permis de découvrir réellement le monde de l'entreprise.

Au final ce travail nous a aidé à développer notre autonomie et nous a permis de mettre à profit la perfection de nos compétences et nos connaissances acquises durant notre parcours académique et d'en acquérir de nouvelles.

Bibliographie

- [1] : Rapport de projet de fin d'études, « MONITORING APPLICATIF Mise en place d'un outil de supervision, » NIDABRAHIM Youssef - KANGA Dominique Bernard, (2016 - 2017) Université Mohammed V-Rabat.
- [2] : Rapport annuel, Société Générale Algérie, 2016.
- [3] : <http://dictionnaire.sensagent.leparisien.fr/base%20de%20donn%C3%A9es/fr-fr/>
- [4] : Mémoire, M^{elle} CHILI Cherifa, M^{elle} AKKOUCHE Malika, (2015/2016), diplôme master ingénierie des systèmes d'information.
- [5] : Livre « Optimisation des bases de données Mise en œuvre sous Oracle », Laurent Navarro ,2010.
- [6] : <https://licenceiset.files.wordpress.com/2016/09/ch5-bd-2016.pdf>
- [7] : https://www.memoireonline.com/10/17/10125/m_La-mise-en-place-d-une-base-des-donnees-pour-la-gestion-de-la-facturation-des-abonnes-de-la-Regid8.html
- [8] <https://www.techopedia.com/definition/1234/relational-database-rdb>
- [9] : http://lagis-vi.univ-lille1.fr/~lo/ens/commun/bd/bd_cours.pdf
- [10] : <https://www.techniques-ingenieur.fr>
- [11] : www.ibm.com
- [12] : <https://fadace.developpez.com/sbgdcmp/>
- [13] : Magazine, USER GROUPE INFORMIX FRANCE, numéro 1-2014
- [14] : Magazine, GUIDE SHARE FRANCE INFORMIX MAGAZINE, mars 2015
- [15] : levillageinformix.blogspot.com/2017/11/vous-etes-developpeur-javascript-node.html
- [16] : Magazine, GUIDE SHARE FRANCE INFORMIX MAGAZINE, juin 2015
- [17] : <http://supervision.clever.fr/monitoring-protocoles-reseaux/>
- [18] : Rapport de stage « Mise en place d'une solution de surveillance permettant de superviser les bases de données », SEUS Max Bruno, 2014, Université de la Réunion.
- [19] : https://www.memoireonline.com/04/12/5604/m_Monitoring-dune-infrastructure-informatique-sur-base-doutils-libres16.html

[20] : <http://www.appperfect.com/services/monitoring/database-monitoring.php>

[21] : <https://www.datadoghq.com/blog/monitoring-101-collecting-data/>

[22] : <https://www.techopedia.com/definition/30560/database-performance-monitoring>

[23]:<https://www.quora.com/What-are-the-various-monitoring-tools-to-monitor-database>

[24] : Rapport de projet de fin d'études, « Mise en place d'un outil de supervision de réseau d'entreprise », Ghada Ben Sassi, (2014/2015), université virtuelle Tunis

[25] : www.centreon.com

[26] : <https://docs.influxdata.com/influxdb/v1.5/>

[27] : ThomasDELANNOY, «Telegraf, InfluxDB et Grafana »,29/09/2017,

[28] : docs.grafana.org/

[29] : <https://whatis.techtarget.com/definition/Perl>

[30] : <https://www.webopedia.com/TERM/C/CPAN.html>

[31] : <https://metacpan.org/pod/DBI>

Résumé

L'objectif principale de notre travail est de réaliser une plateforme de supervision de bases de données au sein de Société Générale Algérie.

Cette plateforme va permettre aux administrateurs de bases de données de surveiller en temps réel les performances des bases de données Informix pour cela nous avons procédé par étapes dans un premier temps nous avons collecté toutes les informations sur l'activité de la base de données Informix, ensuite nous avons utilisé une base de données chronologique pour les stocker. Nous avons aussi utilisé un outil de visualisation Grafana pour la mise en forme et la génération de graphes à partir des données récoltées.

Enfin la plateforme offre un système d'alerting très performant qui permet d'être alerter en cas d'anomalies ou d'incidents au niveau de la base de données et permet d'être notifié par la suite.