

Les types de data :

Mysql, csv, dumpsql, rest, apisoap, apirest, paradox

Userscases :

-Marketing (par quels sites, par quels biais les clients ou les prospects sont arrivé ?)

-Commerciaux (rapport au chiffre d'affaire)

-Chef de projet (Compte rendu d'activité des collaborateurs)

Les applications peuvent-elles être accessibles hors de l'entreprise ?

Oui, en fonction de l'orientation du projet, nous aurons accès à certaines données, par le biais de VPN si nécessaire

Le datawarehouse actuel :

Infrastructure classique, début du projet par les outils : CASSANDRA pour le reporting, ETL avec talent, puis Hadoop mais les trois donnaient quelque chose de couteux et de compliqué.

Données à lier pour le projet : financière, intranet, CRM

Le projet :

Base d'étude

Création d'une infrastructure

Construction de base de données (cassandra, hadoop, mongo)

Travail en lien avec les personnes chargé du développement du datawarehouse

Le déroulement du projet :

1 mois : Etude des plateforme et infrastructure

1 à 2 mois : Mise en place, et test mini poc (ex talent + cassandra ou soap), liste de connecteurs

Vers le 3^{ième} mois, commencement de l'implémentation.

But :

Concurrencer le datawarehouse existant, plateformes et performance

Solution de complémentarité avec le datawarehouse existant si possible remplacer l'existant dans le contexte de Big Data pour concurrencer celui actuel et à l'avenir faire une solution fonctionnelle censé la remplacer

Réussite pour l'entreprise : développer 5 connecteurs types

Données : logs, analyses (risque d'intrusion, historique d'audit)

Finalité, notre profil peut être retenu.