

Journal de bord d'alternance



DEVELOPPONS ENSEMBLE
L'ESPRIT
D'EQUIPE  SOCIÉTÉ
GÉNÉRALE

– Journal n 5 –

Loïc DIVAD

25 août 2016

1er Semestre

1er Septembre

Arrivé et présentation de l'équipe. Présentation des différents projets et de l'écosystème dans lequel travail la petite équipe constituée de mon manager et son chef. Six petits projets sont en cours. Ce sont des POC pour « proof of concept », destinés à évoluer pour devenir de vrai sujets d'industrialisation qui viendront enrichir le porte folio du pôle de compétence Big Data de notre entité.

2 Septembre

J'assiste à la première réunion sur les POC, « proof of concept ». Les technologies utilisées dans ces projets pourront donc servir à d'autre projets et pour différentes entités.

3 Septembre - 5 Septembre

La fin de la semaine à été dédiée à la préparation de mon poste de travail, de mon badge, des mots de passes et d'autres sujets administratifs. Je rencontre les consultants externes qui sont sur les projets depuis le lancement de l'initiative Big Data lors des réunions d'avancement pour les POC.

8 - 12 Septembre

Une des finalités des projets big data est la constitution d'un « data lake » pour notre entité. Nous serons donc amenés à nous intéresser au logiciel Hadoop. Dans cette optique j'ai reçu une documentation pour débiter. Les 800 pages de doc ont été fournies à mon manager à la suite d'une formation. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

15 - 19 Septembre

Cette semaine on m'a présenté les 4 machines dédiées aux POCs en interne. (Ils sont actuellement chez le prestataire CGI). Je m'adapte donc aux machines virtuelles et ce qui est mis à disposition dessus. Je commence l'installation du logiciel hadoop avec la distribution « Apache » sur un seul nœud. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

19 Septembre

Une réunion particulière a également eu lieu. Elle réunit les commanditaires du projet « 360 » (métier, sponsor et maitrise d'ouvrage), notre équipe et les consultants CGI pour la présentation des avancés et des discussions sur les budgets. Elle était particulièrement enrichissante. Elle m'a permis de mieux comprendre dans quelles conditions nous travaillons.

22 - 26 Septembre

Je poursuis l'installation du logiciel Hadoop tout en continuant de progresser dans la documentation. Il y a actuellement des discussions sur l'architecture technique qui concerne en partie la distribution utilisée. Il se peut que je refasse l'installation avec une autre version. Comme il y a un fort aspect veille technologique dans notre travail il peut être intéressant de comparer les deux. C'est un plus que je peux apporter à l'équipe qui n'a pas forcément le temps de le faire

Jeudi 25 Septembre

Rentrée scolaire.

01 - 10 - Octobre

Mon Manageur part en formation pendant une semaine sur les sujets Elasticsearch et hadoop. Je suis donc seul et chargé de « tenir la boutique ». Je suis désigné comme point d'entrée sur le sujet dont il s'occupe notamment la mise en homologation de l'application KYC. Je fais le lien entre les équipes internes et les consultants (sopra et CGI). Il a été difficile de tout comprendre et de se faire comprendre car KYC est le seul projet où je ne suis pas impliqué (ni le développement, ni dans les réunions).

Mercredi 1 Octobre

Premier jours de cours à l'ISEP.

06 - Octobre

Les difficultés rencontrées avec hadoop sont résolues petit à petit. Je prends en note des points bloquants (les dépendances manquantes) pour permettre à mes collègues d'anticiper les prochaines installations. Le principal problème : Les machines de développement pour des raisons de sécurité sont totalement déconnectées d'internet. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

13 - 17 Octobre

Désormais le lancement de hadoop (la distribution d'Apache) se passe sans problème. On m'a indiqué de nouvelles consignes de configuration pour respecter l'arborescence normalisée pas l'entreprise. Je retravaille donc cette partie. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

17 - Octobre

Une réunion très enrichissante à eu lieu aux tours (Défence) avec des participants du secteur bancaire. Il s'agit du point de validation des Use Case pour

le lancement du POC TRACFIN. C'est le POC le plus important dans le sens où c'est celui qui fait intervenir la plus grande volumétrie et des compétences de Data scientist tel que le Machine Learning et la Data Discovery. Les aspects « business » et juridique sont fortement présent puisqu'il s'agit de détection de comportements frauduleux.

20 - 24 Octobre

Un passage de connaissance est assuré par les équipes CGI. Nous nous déplaçons donc mon manager et moi pour assister à une introduction sur les technologies qui composent les petits projets 360. La séance se termine par le passage des sources. Je suis chargé de m'en imprégner. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

27 - 31 Octobre

Vacances de toussaint

Mon manager part en vacance pour la semaine et je suis de nouveau le point d'entrée en cas problème. Je découvre quelques processus société générale. Je termine mes recherches sur la configuration d'hadoop et recommence avec une nouvelle distribution (Cloudera). Je termine et tien au courant mon manager de cette avancée par mail.

Je suis également chargé, à l'aide de mes collègues sur place, de monter 3 réunions dédiées à la recherche de nouveaux cas d'utilisations du projet LogSyris. Je réussit à contacter 5 collaborateurs.

3 - 7 Novembre

La série de trois réunions organisées pendant les vacances débute cette semaine. Elle réunit pour la première fois un nombre étendu de collaborateurs pour une réunion de travail. Je réajuste les ordres de réunions au fur et à mesure et je prépare les comptes rendus. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

10 - 14 Novembre

Mon travail sur hadoop s'arrête ici. Il est déterminé en réunion d'architecture que les outils autour de l'écosystème hadoop seront apportés par une autre entité. Nous nous serviront donc de leur socle technique pour travailler et ne nous soucieront donc pas des installations et configuration. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

17 - 21 Novembre

Participation lundi et mardi à un événement interne société générale : Hackathon SG, un concours de développement.

24 - 28 Novembre

Arrivée de nouvelles ressources pour le lancement d'un nouveau projet. Nous recevons également du nouveau matériel qui leur est dédié. Je suis donc chargé de réaliser quelques installations standards et les accompagner sur ce nouvel environnement de travail. Ils m'expliquent ensuite la dimension « Data analyse » qui est un aspect nouveau dans nos projets. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

01 - 12 Décembre

J'assiste les ressources dédiées au projet TRACFIN et essaie de monter en compétence sur hadoop.

15 - 19 Décembre

Je profite du calme provoqué par les départs en vacances pour prendre de l'avance et rédiger un TP proposé par l'ISEP aux A3, qui portent sur des sujets étudiés à la société générale. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

22 - 26 Décembre

Mon manager étant en vacance il laisse des objectifs à réaliser pour les deux semaines. Avec l'aide de mes collaborateurs externes je dois réaliser la migration de 4 projets vers les environnements sécurisés de la société générale (actuellement sur les serveurs de CGI). Ce qui comprend l'installation des logiciels et le déploiement des sources.

25 Décembre

Joyeux Noël !

29 Décembre - 2 Janvier

Après quelques prises d'initiative la migration d'un des projets prend du retard mais cela m'a permis de respecter des processus standards et d'établir directement une documentation. Tout est installé à temps et des présentations auront lieu dans les semaines qui suivent.

1er Janvier

Bonne Année.

05 - 09 Janvier

Il m'est demandé de m'imprégner des développements réalisés sur un projet en particulier : LogSyris. Un projet d'analyse de l'utilisation d'une application SG. L'objectif étant de pouvoir le décliner pour une autre application. C'est alors un sujet que je pourrai suivre de bout en bout, autant dans les spécifications que dans les développements.

La semaine est donc dédiée à l'analyse du code de ce projet. Il fait intervenir des langages/technologies qui me sont inconnus.

12 - 16 Janvier

Je continue d'analyser et d'assimiler les développements du projet LogSyris. J'en profite pour reformuler et commenter le code qui a été fourni dans un mode POC, c'est à dire en une série de « sprint » très courts.

La semaine est également partagée avec un autre sujet très urgent : la mise en homologation de l'EDD (entrepôt de données). Il m'est donc demandé de rassembler tous ce que j'ai réalisé pendant les vacances de Noël pour l'environnement de développement et de rédiger un document de mise en homologation.

19 - 23 Janvier

Des démonstrations sur le projet LogSyris ont lieu. J'intègre de nouvelles données pour qu'il y ait plus de matière à observer. Je participe aux réunions avec la MO de ce projet.

L'homologation de l'EDD prend du retard au niveau de la mise à disposition des machines. J'en profite et je contacte les consultants d'une SS2i (Théodo) employée par une autre entité pour m'aider au sujet de la cohabitation entre leur application et notre EDD. Je propose de faire des essais dans l'environnement de développement.

19 - Janvier

Mon équipe se réunit pour mettre en place la première offre de service du pôle de compétence de notre entité. Il s'agit de répliquer LogSyris, et de proposer l'analyse de log à d'autres entités. Nous sommes aidés par le manager de l'équipe CGI pour la définition de cette offre de service. Ma première mission en solo sera certainement de répondre à un client SG demandeur de ce type de service. La semaine est également ponctuée par les réunions d'avancement et d'architecture.

26 - 30 Janvier

Je modifie et peaufine le dossier de mise en homologation de l'entrepôt de données (EDD). Je devrais, normalement la semaine prochaine, piloter la livraison et faire le lien avec les opérationnels qui déploient l'application.

La dernière intégration de données se passe très mal pour le projet logSyris. Je prend en note les améliorations que j'aimerais apporter et je les expose.

2ème Semestre

02 - 06 Février

Je suis chargé retrouver l'origine d'un bug dans l'application 360. Cette application fait l'objet de plusieurs présentations cette semaine. Le recule pris sur les sources me permet de corriger l'erreur en une demi journée. Le reste de la semaine est consacré à l'avancement sur le projet Tracfin.

09 - 13 Février

Je travaille cette semaine sur les scripts d'intégration de l'application Log-Syris. J'automatise l'intégration des logs, je documente les fonctions et rajoute quelques commentaires. Jusqu'à maintenant l'intégration des sources de données était presque manuelle. Même si il n'est pas encore décidé d'industrialiser le projet je met en place solution intermédiaire pour faciliter le chargement.

16 - 20 Février

Développement de scripts qui pourraient être utiles pour l'intégration des fichiers que l'on est amené à traiter.

23 - 27 Février

Les deux jours de cette semaine sont consacrés à la livraison du projet EDD (entrepôt de données) dans l'environnement d'hologation.

02 - 06 Mars

Livraison en Homologation EDD n1.

09 - 13 Mars

Cette semaine j'assiste de nouvelles équipes dans leur installation et prise en main des environnements de développement. En effet les prestataires se renouvellent et parmi les nouveaux arrivants se trouvent des consultants ORACLE. Ce qui va certainement nous rapprocher de technologie propriétaires.

16 - 20 Mars

Cette semaine je prépare la mise en homologation n 2 de l'entrepôt de données (EDD). Je prépare une documentation pour les opérateurs en prenant en compte les problèmes rencontrés sur la livraison du premier lot. Je retire également les parties qui ne sont pas à répéter. La semaine comprend aussi des corrections de bugs sur l'application "parcours utilisateur".

23 - 27 Mars

Livraison en Homologation EDD n 2.

30 Mars - 03 Avril

Dans le cadres du projet 360 des ouvertures de routes sont réalisées. Une erreur à été comise lors de l'ouverture et la mauvaise machine a été traitée. Je dois donc échanger deux applications de machine et les remettre en route

06 - 10 Avril

L'un des pocs de notre entité est décliné en offre de service et un premier client (d'une entité voisine) nous propose un point pour savoir si l'ont peut apporter un valeur intéressante pour sont application. J'ai donc les logs de cette application et une journée pour en extraire tous ce que je trouve et le présenter.

Avec un script python je crée quelques graphiques que je pose sur un slide accompagnré de plusieurs métrique repérée. Les notions apporté par le cours de System d'exploitation ont été très utiles pour cette tâche.

13 - 17 Avril

Je relance l'application Parcours utilisateur qui à cessé de fonctionné en début de semaine. Une dixième réunion à lieu avec un client potentiel à l'offre de service Parcours utilisateur.

20 Avril

Vacances de Pâque.

20 - 24 Avril

La grande mission des vacances est d'utiliser les ouvertures de route établir les flux entre nos petites applications et divers site d'information / réseaux sociaux.

04 Mai - 29 Juin

A la suite d'une réorganisation du service il est déscidé que les projets de type suivit de parcours utilisateurs sont mis de coté. Je me concentre donc sur les projets 360 qui vont repartirent pour de nouveaux cycle de développements dans les environnement SG pour de nouveaux commanditaires.

Juillet - Août

Après la fin des cours, le 29 Juin, débute une période continue en entreprise. Je suis amené à revenir sur certains points qui sont plus difficiles à suivre à mis-temps.

- Les décisions de sécurités et d'architecture liées à l'initiative du "Data-lake"
- La nouvelle architecture mise en place dans le cadre de la reprise du POC 360.

Dans le cadre du "DataLake" il s'agit essentiellement d'une prise de connaissance sur le socle technologique mis en place. En effet les prochains POC ou projets pourraient cohabiter avec le "DataLake". La deuxième tâche a un aspect bien moins théorique puisqu'il me faut monter en compétence sur les technologies apportées (par les prestataires OCTO) dans le cadre de la version deux du "POC vision 360". La cible pour moi étant de participer activement au développement d'un projet.

Dès le 25 juillet j'agis sans mon supérieur qui est en congé. Durant cette période je suis principalement chargé d'assurer un suivi sur le POC EDD "Entrepôt de donnée". Le projet est relancé par l'unique abonné au service délivré par l'EDD (le Dashboard Interactif). Le mois est donc consacré à la correction des bugs constatés par nos collaborateurs de l'entité COO, responsable du Dashboard et l'établissement des différents points liés à l'industrialisation de l'EDD. En dehors de développements, J'interviens également sur l'enrichissement des spécifications techniques et la documentation.

10 - 25 Septembre

Dès la rentrée, une nouvelle demande provenant des équipes COO remet en question le travail effectué sur le projet EDD lors des mois Juillet- Août. Il s'agit d'historier des informations transmises aux équipes FAST-IT (en charge du Dashboard interactif RESG). Un collaborateur (FAST-IT) dédié au problème m'aide à comprendre ce qui est attendu. Nous parvenons donc au résultat à la fin du mois. Le premier mois du troisième semestre se partage donc entre cette tâche et la préparation de ma soutenance de mi-parcours.

3ème Semestre

01 Octobre - 29 Janvier

La fin de l'année est principalement dédiée à l'entrepôt de données (l'EDD). Les actions d'amélioration sont limitées au profit de l'exploitation de l'application. J'apporte mon soutien lors des réunions sur la partie de l'application traitée par notre entité (RESG/BSC/DAT). Une ressource supplémentaire est désignée pour m'aider à assurer l'exploitation de cette application. J'ai donc l'occasion de transmettre mes connaissances, en plus de la documentation écrite par mes soins sur le sujet.

Je suis également sollicité pour rédiger des scripts d'aide à l'intégration de données pour l'application KYC. Ces scripts sont liés à des besoins utilisateurs révélés lors des différentes itérations du cycle AGILE.

4ème Semestre

01 Février - 29 Juillet

Lors du dernier semestre d'alternance les clients du projet DBI décident d'arrêter l'application du Dashboard interactif RESG. Elle est reliée au POC EDD qui se trouvait déjà en standby. Cela signe la fin du POC EDD. J'arrête alors d'assurer la production de l'application. L'environnement de développement sera bientôt réassigné. L'application BCC (anciennement KYC) elle aussi est arrêtée. Les directeurs du programme Big Data ont estimé qu'elles se rapprochaient trop de l'application CBS. Une migration sera donc bientôt planifiée, de sorte à ce que CBS couvre entièrement le périmètre de BCC (en terme d'informations accumulées).

Je termine mon alternance avec une nouvelle équipe développeur. L'équipe du projet Maille Contrat. J'ai pour tâche d'implémenté les fonctionnalités demandées par les utilisateurs (user story) du projet. Je participe aux réunions de définition des fonctionnalités et de démonstration. J'interviens donc sur le chiffrage en jours de quelques modules. Ce projet est l'occasion de confirmer mes compétences en Scala et en Spark.