
Les exigences pour un choix d'architecture dans le cadre d'une migration dans les nuages

Antoine Aubé¹, Thomas Polacsek², Clément Duffau³

1. *Office National d'Études et de Recherches Aérospatiales*
Toulouse, France

antoine.aube@onera.fr

2. *Office National d'Études et de Recherches Aérospatiales*
Toulouse, France

thomas.polacsek@onera.fr

3. *Stack Labs*

Toulouse, France

clement.duffau@stack-labs.com

RÉSUMÉ. La migration d'un système vers un nuage est un ensemble d'activités permettant de tirer profit de l'informatique en nuage. Elle nécessite la conception d'un environnement d'exécution déployé par la configuration de services infonuagiques. Du choix de ces services et de leur configuration dépend la satisfaction de certaines exigences qui concernent le développement et l'exploitation du système à migrer. À travers une série d'entretiens semi-dirigés auprès de spécialistes de l'informatique en nuage, nous avons identifié dix exigences de haut niveau pour la sélection d'un environnement infonuagique lors d'une migration et leurs conséquences sur la conception de systèmes, ainsi qu'une classification de ces exigences.

ABSTRACT. A cloud migration is a set of activities to take advantage of cloud computing. It requires the design of an execution environment deployed by configuring cloud services. The satisfaction of certain development and operations-related requirements depends on the choice of these services and their configuration. Through a series of semi-structured interviews with cloud computing specialists, we identified ten high-level requirements for selecting a cloud environment during a migration and their implications for system design, together with a classification of these requirements.

MOTS-CLÉS : *Informatique en nuage, Migration, Ingénierie des exigences*

KEYWORDS: *Cloud Computing, Cloud migration, Requirements engineering*

1. Introduction

L'informatique en nuage, dit *Cloud Computing* en anglais, est un mode d'accès à des ressources informatiques comme des serveurs, du stockage, du réseau voire du logiciel, sous la forme de services accessibles via l'Internet. Une telle approche repose sur une infrastructure matérielle mise en commun pour tous les utilisateurs qui peuvent individuellement allouer et désallouer très rapidement des ressources suivant leurs besoins (Mell, Grance, 2011). Ainsi, en choisissant et en configurant ces ressources, l'utilisateur constitue un environnement sur lequel il peut déployer son application. Avec l'informatique en nuage, la notion d'infrastructure tend à s'effacer. Dans les faits, elle est masquée par un catalogue de services nommé un nuage. Les catégories communément utilisées pour classer ces services définissent les responsabilités portées par le fournisseur infonuagique et ses utilisateurs. Ces catégories sont, ordonnées par responsabilité croissante du fournisseur infonuagique : *IaaS*¹, *PaaS*² et *SaaS*³.

Pour une entreprise, il peut être intéressant en termes de coûts de passer à l'informatique en nuage. En effet, de nombreuses entreprises proposent aujourd'hui des services infonuagiques, ce qui permet aux clients d'uniquement payer les services à l'usage, c'est-à-dire ce qu'ils consomment réellement. Dès lors, l'entreprise peut espérer réaliser des économies, puisqu'elle n'a plus à constituer et entretenir un parc de machines avec tous les coûts inhérents tels que le loyer, l'électricité, l'accès à l'Internet, etc. De plus, sans machine, elle s'exonère des tâches de maintenance et du personnel nécessaire à la maintenance.

Cependant, la réalité n'est pas si simple. On ne conçoit pas et on ne développe pas des systèmes infonuagiques comme on conçoit un système d'information classique. En effet, pour qu'un système infonuagique soit efficace, celui-ci doit utiliser les services adéquats, sans surconsommation tout en garantissant son exécution opérationnelle. Nous avons ici un double objectif qui, d'une part, est de définir et provisionner les bons services pour répondre au besoin du système et, d'autre part, de ne pas surprovisionner pour éviter des coûts inutiles.

Cet objectif est très difficile à gérer dans les phases en amont d'un cycle de développement. En effet, il n'est pas simple d'évaluer, en phase de spécification, les effets de l'utilisation de service sur des exigences portant, par exemple, sur les coûts ou la qualité du système. Dans la pratique, ce manque d'anticipation est la cause de dépassement de budget et de systèmes qui ne répondent pas tout à fait à leurs

¹1. *IaaS* pour *Infrastructure as a Service*, service d'infrastructure à la demande : le service déploie des unités de calcul, du stockage et de la mise en réseau.

²2. *PaaS* pour *Platform as a Service*, service de plateforme à la demande : le service déploie un cadre d'exécution des logiciels et prend en charge la gestion du système d'exploitation sous-jacent. Dans la pratique, le client a le contrôle sur l'application déployée, mais pas sur l'infrastructure virtuelle.

³3. *SaaS* pour *Software as a Service*, service de logiciel à la demande : le service déploie un logiciel prêt à l'usage (e.g. Gmail, Gitlab.com).

exigences. Il est donc nécessaire de disposer de la capacité à évaluer, le plus tôt possible, les conséquences des choix de services sur les exigences du système. Encore faut-il comprendre les motivations qui mènent aux choix de services et aux choix du fournisseur infonuagique. Dès lors, même si ces motivations sont multiples, il peut être intéressant de dresser un panorama des éléments qui impactent fortement le choix du fournisseur et des services. Par éléments, nous entendons des exigences de très haut niveau, des buts et des objectifs. Ces buts peuvent être fonctionnels, comme garantir une performance pour un nombre de requêtes donné, ou non-fonctionnels, comme disposer d'un hébergement dans une zone géographique bien précise.

Pour pouvoir évaluer et ainsi fournir une aide à la conception de systèmes infonuagiques, il nous semble important de cartographier l'ensemble des dimensions qui sont cruciales pour les concepteurs de tels systèmes. C'est suivant ces buts, ces dimensions, qu'il sera possible d'évaluer des choix d'architectures et de faire des compromis entre plusieurs solutions.

Dans cette optique, nous avons mené une première enquête, au travers d'entretiens avec des spécialistes de la migration d'applications en informatique en nuage. La migration vers un nuage consiste tout simplement à faire migrer des applications d'un système d'information local à l'entreprise vers une solution utilisant de l'informatique en nuage. Dès lors, une telle migration implique des changements d'architecture et des choix sur les services qui vont être utilisés. De plus vient le problème du choix des fournisseurs de ces services. Nous allons ici présenter les premiers résultats de cette enquête.

Dans la Section 2, nous présentons l'objectif de notre recherche et la méthodologie que nous suivons. La Section 3 décrit l'échantillon de participants à l'enquête et la Section 5, suivi d'une proposition de classification de ces exigences en Section 6. Enfin, nous présentons des perspectives pour des travaux futurs en Section 7 avant de conclure en Section 8.

2. Objectif, motivations et méthodologie

Notre objectif est de comprendre ce qui motive les choix d'environnements infonuagiques dans une opération de migration. Par environnement, nous parlons des ressources déployées par la configuration de services infonuagiques. Nous avons choisi de nous limiter ici à la migration afin de ne pas étendre le champ des possibles à tous les projets de systèmes infonuagiques. Ce cadre permettant, nous l'espérons, de dégager un premier corpus restreint d'exigences. Nous pouvons ramener notre but à la question suivante :

Quelles sont les principales exigences de la sélection d'un environnement dans le cadre de la migration vers un nuage ?

Pour pouvoir répondre à cette question, nous avons décidé de mener des entretiens de spécialistes du domaine. Notre approche est motivée par le manque de données dans la littérature. En effet, si de nombreux travaux s'intéressent à identifier

des stratégies de choix d'environnement, le sujet des motivations des organisations qui dirigent une migration vers un nuage est peu adressé.

Citons, pour exemple, les travaux de Quinton (2014) qui propose une méthode basée sur les lignes de produits logiciels pour sélectionner un environnement à partir de l'expression des besoins du système. Dans le même esprit, Ferry *et al.* (2013) proposent un langage de modélisation d'infrastructure afin de faciliter l'utilisation de ressources infonuagiques. L'utilisateur exprime les besoins non fonctionnels de son système dans un modèle indépendant de tout fournisseur infonuagique, puis ce modèle est instancié pour un fournisseur donné et utilisé à l'exécution pour le déploiement, la surveillance et l'approvisionnement dynamique des ressources. Toujours dans une logique d'aide à la sélection de service, Zhang *et al.* (2012) ont conçu un système de recommandation pour trouver la meilleure configuration de machines virtuelles en fonction d'exigences non fonctionnelles. Leur approche se base sur l'ontologie des ressources de l'informatique en nuage : *Cloud Computing Ontology* (CoCoOn). Pour finir, Amiri *et al.* (2017) proposent d'utiliser l'apprentissage automatique pour optimiser le provisionnement des ressources.

Comme nous le voyons, tous ces travaux adressent la satisfaction des exigences, mais ne se focalisent pas sur leur identification ni sur les motivations sous-jacentes.

Forts de ce constat, nous avons opté pour une approche basée sur un retour d'expérience. Plus précisément, nous avons mené une série d'entretiens individuels qualitatifs semi-structurés auprès de professionnels. L'entretien qualitatif est une pratique courante et assez ancienne dans le champ de l'ethnographie et de la recherche médicale (DiCicco-Bloom, Crabtree, 2006). Il permet au travers d'une conversation approfondie, suivant un protocole préétabli, d'acquérir une meilleure compréhension du sujet d'étude et de formuler des hypothèses. Généralement semi-structurés, de tels entretiens consistent à poser des questions ouvertes prédéterminées.

Notre méthodologie a été la suivante. Nous avons mené une série d'entretiens individuels. Ces entretiens ont été menés en face à face, ou en visioconférence, sur une durée allant de 45 minutes à une heure. Pour faciliter la retranscription, ils ont été enregistrés avec l'accord des participants.

Les entretiens ont toujours suivi le même format avec, pour commencer, une phase d'introduction avec une explication du sujet de recherche, une explication du périmètre de la discussion afin de limiter les digressions sur des sujets connexes, et une présentation du déroulement de l'entretien.

L'entretien étant semi-structuré nous avons défini un jeu de questions préétablies posées systématiquement et toujours dans le même ordre. La première série de questions se rapporte au participant, ses rôles au sein de l'équipe, et au contexte de la migration (taille de l'équipe de développement, domaine de l'entreprise, etc.). Ensuite commencent réellement les questions concernant notre sujet d'étude. Nous avons dérivé notre question de recherche sous la forme des dix questions ouvertes ci-dessous :

1. Comment avez-vous analysé les composants à migrer ? Quelles étaient les informations importantes à en tirer ?
2. Quelle a été la démarche d'analyse pour trouver les besoins spécifiques à l'informatique à nuage ? Autrement dit, pour identifier les exigences concernant la sécurité, la qualité de service, etc.
3. Sous quelle forme avez-vous pu consulter les exigences auxquelles répond le système ?
4. Avez-vous des preuves que le système existant répondait effectivement aux exigences ?
5. La façon dont l'utilisateur se sert du système était-elle un élément important à capturer à cette étape ? Si oui, comment cela a-t-il été fait ?
6. Avez-vous des exigences sur la qualité de service ? Si oui, quels aspects de la qualité de service ?
7. Quand on parle budget opérationnel, est-ce que le coût de la main-d'œuvre est pris en compte ? Et si oui, comment est-il calculé ?
8. Avez-vous des exigences sur la performance environnementale du système ?
9. Pour le budget opérationnel, comment sont exprimées les exigences ? Est-ce un budget global au système, un par unité fonctionnelle, par ressource, par volumétrie d'entrée, etc. ? Est-ce que ce sont des sommes fixes ou bien est-ce plus complexe que ça ? Est-ce que nous parlons de seuils de tolérance, de marge, etc. ?
10. L'effort d'adaptation du code a-t-il été une contrainte ?

3. Profils des interlocuteurs

Afin d'obtenir des réponses pertinentes lors des entretiens, McCracken (1988) recommande la constitution d'un échantillon homogène ayant une proximité avec la question de recherche. Pour cette raison, nous avons choisi de nous focaliser sur des personnes proches des aspects techniques d'une migration vers un nuage et qui sont intervenus dans les mêmes conditions. Pour mener notre enquête, nous avons interrogé sept professionnels qui travaillent dans des sociétés de services numériques extérieures aux organisations qui possèdent le système à migrer. Concernant la proximité avec notre objet d'étude, tous ces professionnels occupent un poste en lien avec le choix de l'architecture de l'environnement infonuagique, que cela soit en terme de conseils, d'étude ou de décisions techniques. Ainsi nous avons interrogé un responsable projets, trois architectes de systèmes infonuagiques, deux spécialistes de l'exploitation opérationnelle et un expert de la migration, qui a plus d'une cinquantaine de migrations à son actif.

Sans prétendre à l'exhaustivité, nous avons cherché des personnes qui ont participé à des migrations très différentes, aussi bien dans la taille des projets et leurs objectifs que dans leurs contextes. En termes de contexte, les projets couverts par nos entretiens adressaient : la migration de systèmes pour une entreprise qui gère

de très gros volumes de données avec des clients dans le monde entier, des migrations pour un établissement public, une migration pour un service à destination des collectivités locales et des migrations de plusieurs centaines de systèmes pour un grand groupe de distribution. Concernant la taille des projets, certains systèmes étaient maintenus par une équipe très réduite avec un développeur commun à plusieurs systèmes qui intervient également sur l'exploitation des systèmes, un ou deux experts du métier qui sont aussi développeurs et un chef de projet. De l'autre côté du spectre, nous avons aussi pu discuter de migrations de systèmes maintenus par des équipes de plus de cinquante personnes.

4. Entretiens

Une première remarque préliminaire s'impose concernant le contenu des entretiens. À notre grande surprise, il est apparu que les développeurs ne connaissaient pas le terme *exigence* et que la définition d'exigences n'était pas toujours exsangue de certaines ambiguïtés pour les architectes. Cependant, au vu des entretiens, nous sommes bien dans l'identification des objectifs, des buts et des contraintes du client qui mènent à définir l'environnement infonuagique, nous avons bien des interlocuteurs qui nous parlent d'exigences de haut niveau. Par conséquent, dans un souci de clarté, nous utiliserons le mot « *exigence* » dans la suite, même si ce mot n'a pas été employé par les personnes interrogées.

Par ailleurs, lors des entretiens, il est important de noter que nos différents interlocuteurs ont relaté des difficultés à accéder à certaines exigences. Par exemple, dans le cadre d'un projet consistant à identifier les problèmes de systèmes existants et à proposer un environnement infonuagique pour une migration, il a été fait état d'une non-connaissance du budget opérationnel du futur système : « *Je pense qu'ils l'avaient en interne, mais à aucun moment ils nous l'ont dit ; je pense qu'ils attendaient [notre proposition] pour voir si c'était une solution qui irait pour leur besoin.* ». Sur ce point, nous ne pouvons faire que des supputations. Il est possible que l'information concernant le budget n'ait pas été communiquée dû à la relation commerciale entre l'entreprise et l'intervenant, ou parce que l'entreprise, manquant de connaissance sur l'informatique en nuage, n'était pas en mesure d'estimer ce qu'elle était prête à payer pour l'opération du système.

Un élément important, présent dans tous nos entretiens et dans tous les projets auxquels ont participé tous nos interlocuteurs, est la présence d'exigences liées aux coûts. Ceci n'est pas étonnant. En effet, la réduction des coûts est la promesse faite aux organisations pour promouvoir l'informatique en nuage. Dès lors, il est normal que les coûts soient un aspect fondamental dans la sélection de l'environnement. Nous devons toutefois faire ici le tri parmi les coûts possibles. Premièrement, nous avons les coûts liés aux activités de la migration. Ces coûts adressent tous les développements logiciels éventuellement réalisés pour adapter le système ainsi que toutes les opérations de déploiement sur le nuage. Deuxièmement, les coûts opérationnels. Ce sont des coûts récurrents et ils concernent le l'exploitation du système.

Concernant les coûts opérationnels, nous pouvons aussi les diviser en deux catégories : les coûts de l'environnement infonuagique et les coûts liés à la maintenance opérationnelle. Les coûts de l'environnement infonuagique correspondent tout simplement à ce que facture le fournisseur infonuagique par l'utilisation de ses services. Ils sont parfois estimés à l'avance grâce à des outils édités par ces fournisseurs. Les coûts de maintenance opérationnelle correspondent, eux, au personnel de l'organisation chargé d'exploiter le système, et ils sont très difficiles à anticiper.

5. Exigences motivant les choix d'architecture

À partir des entretiens, nous avons identifié dix catégories d'exigences, d'exigences de haut niveau, motivant les choix lors de la constitution d'un environnement infonuagique dans le cadre d'une migration. Elles sont présentées dans cette section.

Notons que ces exigences ne sont pas apparues nommément dans le discours des personnes interrogées. Leur identification relève de notre part d'une généralisation d'un ensemble d'anecdotes et d'exemples d'exigences et de contraintes très factuelles données par les personnes interrogées. Par exemple, une anecdote collectée relevait d'une contrainte sur l'utilisation d'un langage de programmation : *« Il fallait que ça soit maintenable par eux en termes de développement. [...] Par exemple, les scripts en Python, on ne pouvait pas proposer qu'ils soient réécrits en Go, car ils ont des ingénieurs qui viennent du monde de l'intelligence artificielle, qui font du Python, car c'est très répandu dans leur domaine, et c'est difficile de leur faire changer de langage du jour au lendemain »*. En prenant du recul, nous constatons qu'il s'agit d'un cas particulier d'un critère plus général : minimiser le besoin de nouvelles compétences pour développer et opérer le système.

Notre étude étant limitée à un échantillon très restreint, il serait illusoire de chercher à déterminer une quelconque relation de préférence entre ces exigences. Dans le cadre de cet article, nous avons donc décidé de présenter ces dix exigences en les classant en ordre décroissant suivant le nombre d'interlocuteurs qui les ont évoquées.

5.1. Minimiser la maintenance opérationnelle

La maintenance opérationnelle est l'effort fourni par des opérateurs humains pour entretenir le bon fonctionnement du système et assurer la livraison dans de bonnes conditions du service.

La diminution de cet effort est apparue dans l'ensemble de nos entretiens comme un critère de sélection de services infonuagiques. Au vu des entretiens, il apparaît clairement que le but sous-jacent à cette exigence est la diminution des coûts de la masse salariale dévolue à l'exploitation du système. Selon le contexte, elle ne s'instancie pas de la même façon, car, dans certains projets, elle vise à réduire la

charge de travail des équipes déjà en place là où, dans d'autres, c'est en prévision du futur pour limiter l'embauche de nouveaux spécialistes.

Comme nous l'a rapporté un de nos interlocuteurs, cette exigence peut amener à une adaptation de l'architecture de migration. Dans son cas, il citait l'exemple de l'automatisation de tâches manuelles en utilisant un service de *Apache Airflow*⁴ à la demande.

Dans le cadre de projets visant à proposer des environnements infonuagiques pour des systèmes de traitement de très grosses données, les architectures privilégiaient systématiquement l'utilisation de services *serverless*⁵ afin de réduire significativement le besoin d'interventions humaines pour assurer la continuité du service.

Dans le cadre de la migration de plusieurs applications impliquant des équipes de maintenance de l'ordre d'une cinquantaine de personnes, tous les systèmes ont été migrés vers un *PaaS* car la maintenance des plateformes monopolisait beaucoup les équipes opérationnelles.

5.2. Minimiser les coûts de l'environnement

Comme évoqué précédemment, un environnement infonuagique est déployé par des services dont les coûts dépendent de l'usage fait du système.

Pour accomplir un même rôle (e.g. stocker des fichiers, exécuter un programme), plusieurs services sont candidats et ne facturent pas le même usage. Par exemple, pour le stockage de fichiers, certains services reposent sur une machine virtuelle dont le coût est facturé (i.e. le temps de CPU et RAM, licence de système d'exploitation, etc.) avec un supplément pour l'infogérance ; d'autres facturent uniquement les interactions (i.e. lectures, écritures, etc.) et le volume de données moyen stocké pendant la période facturée. Le choix de la solution la plus économe est en fait dépendant du modèle d'utilisation du système.

Pour un projet d'hébergement d'une forge logicielle (i.e. une plateforme en ligne qui agrège des outils à l'usage des développeurs pour sauvegarder les codes sources, effectuer des validations, gérer des projets, etc.), en outre des préoccupations quant à la performance, la personne interrogée a décidé d'opter pour un *IaaS* pour pouvoir configurer une instance correspondant le plus possible aux besoins du système. Les *PaaS* qu'il aurait pu envisager lui permettait de configurer des instances soit insuffisantes, soit surdimensionnées. En sélectionnant plus finement la configuration de l'instance, il a ainsi réduit les coûts de l'environnement au détriment de l'effort de maintenance opérationnelle.

Tous les participants ont indiqué que la diminution des coûts de leur environnement était un critère de sélection des services, mais seul un en a fait une priorité. Dans la plupart des cas, cette diminution des coûts de l'environnement

⁴4. Logiciel d'orchestration de tâches de calcul. <https://airflow.apache.org/>

⁵5. Services dans lesquels le fournisseur de services gère entièrement les ressources allouées.

semble moins importante que d'autres critères. Par exemple, dans le cas du projet utilisant *Apache Airflow*, il aurait été envisageable d'employer des services *serverless* coûtant moins cher d'après leur modèle d'utilisation du système, mais cela aurait nécessité des connaissances que n'a pas l'équipe de développement.

5.3. Maximiser la fiabilité du système

Dans le cas d'un hébergement de l'infrastructure informatique sur site, remplacer des ressources peut prendre du temps (e.g. achat, transport, réception, mise en place du matériel) et se protéger d'un incident est coûteux (e.g. réplication des données dans plusieurs centres informatiques suffisamment éloignés).

Quatre de nos interlocuteurs sont intervenus sur des projets pour lesquels la migration a été initiée suite à un tel incident. Ces derniers peuvent être liés, par exemple, à des aléas climatiques, à des actions malveillantes ou à l'usure du matériel, qui ont occasionné des coupures du réseau, la destruction de disques durs ou d'unités de calcul. À cause de ces incidents, leur système n'a pas été utilisable pendant un temps et certaines données ont été perdues.

L'un des objectifs de ces migrations était donc l'amélioration de la fiabilité de ces systèmes.

Un des architectes interrogés rapporte : « *Le but [de la migration] pour l'exploitation de leur système, c'était d'apporter de la robustesse [...] Le but premier, c'était de fiabiliser, car, avant, tout reposait sur quelques machines et un ingénieur qui opérait l'ensemble. [...] Avec ce qu'ils avaient avant, s'il y avait une grosse panne, ils auraient eu beaucoup de mal à tout remettre en ordre* ».

Pour ce faire, les risques ont été identifiés et un Plan de Reprise d'Activité (*PRA*) a été mis en place. Ce *PRA* décrit les procédures permettant de réagir à une panne, et repose sur l'utilisation de services qui fournissent des garanties de fonctionnement (*Accord de niveau de service*⁶) par exemple pour la durabilité et la disponibilité des données, ainsi que sur l'automatisation de la réplication des ressources (e.g. données, machines virtuelles) sur plusieurs zones géographiques.

5.4. Maximiser les performances

La performance est un critère que l'on retrouve dans de nombreux projets, mais pas dans tous et ils n'adressent pas tous la même notion de performance.

Plus précisément, dans le contexte d'une entreprise dont l'infrastructure était hébergée en France, l'arrivée d'un nouveau client basé en Amérique du Sud a introduit une exigence de réduction de la latence des services spécifiquement pour ce client. Le choix du fournisseur infonuagique a donc été fait en tenant compte de

⁶ De l'anglais *Service-Level Agreement*, engagement du fournisseur de services sur la qualité du service qu'il délivre.

la possibilité de déployer des ressources à la fois en Europe de l'Ouest et en Amérique du Sud.

Un des objectifs de la migration de la forge logicielle évoquée précédemment était de diminuer les temps d'exécution des tâches de calculs intensifs (e.g. compilation des logiciels). Cette exigence a motivé l'utilisation de services de machines physiques à la demande (sans surcouche de virtualisation coûtant du temps de calcul) dotées de GPU. Cette solution a été préférée par rapport à l'utilisation d'un *SaaS* ou d'un hébergement sur un *PaaS*, qui auraient été des solutions plus simples à gérer. À ce sujet, la personne interrogée nous a dit : « *Il y avait deux choix : soit partir sur du Kubernetes infogéré, soit monter soi-même son cluster Kubernetes. On a fait le second choix pour des raisons de performances. C'est-à-dire que sur les machines virtuelles que propose [le fournisseur infonuagique] derrière leur Kubernetes infogéré, on ne retrouvait pas les caractéristiques dont on avait besoin, ou en tout cas ça aurait eu un coût trop important [...]* ».

5.5. Minimiser le besoin de nouvelles expertises

Travailler avec des services infonuagiques requiert l'acquisition de connaissances techniques très spécifiques (e.g. limitations, protocoles) aussi bien pour le développement que pour l'exploitation. Ces connaissances sont plus ou moins nombreuses suivant les services utilisés. Par exemple, l'administration d'une machine virtuelle est proche de l'administration d'une machine réelle, mais n'est pas comparable avec l'administration d'une ressource *serverless*.

Or, certains projets qui nous ont été présentés ont une équipe de développement réduite avec, par exemple, des développeurs chargés également d'administrer le système. Dans aucun de ces projets, l'entreprise n'envisageait d'agrandir l'équipe, pour des raisons de coûts, ni de remplacer du personnel, par crainte de perdre une expertise métier propre à l'organisation. Le choix des services devait donc minimiser le nombre de nouvelles connaissances à acquérir pour les équipes en place.

Précédemment, nous avons évoqué un retour au sujet d'une migration dans laquelle le langage de développement devait absolument rester *Python*. Cette contrainte a eu de véritables conséquences sur le choix des services infonuagiques : la plupart des *PaaS* prennent en charge en ensemble limité de langages.

Concernant le choix du fournisseur infonuagique, dans le cadre de la migration d'un *SaaS*, un de nos interlocuteurs évoque le fait que l'équipe qui administre le système avait de l'expérience avec un fournisseur et que cela a grandement encouragé le choix de rester avec celui-ci au détriment de services proposés par d'autres.

5.6. *Minimiser les coûts de migration*

Pour bénéficier des avantages de l'informatique en nuage (i.e. élasticité, etc.) ou pour optimiser les coûts de l'environnement infonuagique, il est nécessaire d'adapter les applications. Ces modifications correspondent à du développement logiciel et peuvent considérablement augmenter les coûts de la migration.

Cependant, dans certains projets avec un budget réduit alloué à la migration, il a été décidé de n'apporter que le strict minimum de modifications aux applications, ce qui apporte une contrainte au choix des services capables de les héberger.

En effet, dans certains cas spécifiques, un composant logiciel peut être hébergé sur un *PaaS* sans changement de son code s'il respecte certaines contraintes, comme a remonté l'un des architectes pour un logiciel écrit en Java et exécuté sur un serveur Tomcat qui a été migré sur *Google App Engine*⁷. L'alternative est d'héberger ces applications sur des machines virtuelles, qui permettent une plus grande liberté au prix d'une facturation moins intéressante et d'une moindre infogérance du fournisseur infonuagique.

5.7. *Minimiser la durée de migration*

Certaines migrations sont contraintes par leur durée, avec une échéance de fin fixée en amont qui ne peut pas être changée.

Ainsi, l'un des participants, qui est intervenu dans un projet pour l'agriculture, nous a rapporté « *Comme leur activité est saisonnière, il fallait finir la migration avant la saison qui arrivait* ». Pour tenir les délais, ils n'ont pas effectué un changement de type de base de données qui aurait permis de limiter le coût de l'environnement infonuagique. Également, dans le cadre de la migration d'application de gestion de données massives pour un service international, eu égard au temps alloué pour effectuer la migration, il a été décidé de réutiliser des morceaux de logiciels existant dans le but de ne pas investir de temps et de risque de dérive dans des opérations d'adaptation du système.

Nous sommes donc bien face à une exigence différente de la minimisation du budget de migration, car ici, c'est la date de mise en production qui a dirigé les choix d'architecture.

5.8. *Garantir la souveraineté numérique*

Le concept de souveraineté numérique, qui ne renvoie pas à des contraintes ni des pratiques précises, est de plus en plus prégnant dans le monde de l'informatique en nuage. Des réglementations telles que le *CLOUD Act* pour les États-Unis ou le Règlement Général sur la Protection des Données (*RGPD*) pour l'Union Européenne ne viennent que renforcer cette thématique.

⁷⁷. App Engine, un PaaS de Google Cloud. <https://cloud.google.com/appengine>

Dans la pratique, certaines organisations, en particulier celles qui manipulent des données sensibles pour un État, tendent à prendre en compte localisation des données et des traitements tout comme la législation qui s'applique aux fournisseurs infonuagiques.

Ainsi, cette exigence est apparue lors de projets de migration de deux systèmes pour un établissement public et lors d'une migration d'un *SaaS* à destination des collectivités locales. Concrètement, il était demandé que les données et traitements soient sur le territoire français et que le fournisseur de services soit sous législation française. D'après l'un des spécialistes de l'exploitation opérationnelle : « *Le DSI de [l'entreprise] souhaitait que la solution reste en France. C'était pour du nucléaire français, il n'y avait pas vraiment de raison que ça aille chez [un nuage états-unien] [...]* ». Selon le contexte, il semble que la mise en œuvre satisfaisante de cette exigence puisse prendre des formes variées, car l'expert en migrations a rapporté : « *Ils voulaient que leurs données soient sur un cloud souverain, on leur a proposé un entre-deux : mettre leurs données sur [un nuage états-unien] et la clé de chiffrement sur un cloud français* ».

5.9. Maximiser la sécurité

De façon surprenante, une seule des personnes interrogées a évoqué des exigences liées à la sécurité. Dans son cas, suite à une première étape de migration sur un *IaaS*, des modifications ont été apportées et certains services ont été proscrits. En effet, l'utilisation de ces services impliquait de facto une modification de la topologie du réseau. Cette modification était vue comme un risque de sécurité dans ce contexte. Le système héberge des logiciels que l'État français impose de protéger de tout risque d'espionnage dans le cadre de la *Protection du potentiel scientifique et technique de la Nation*.

De façon générale, les exigences de sécurité peuvent être un frein à l'utilisation de certains services et configurations de services.

5.10. Maximiser l'agnosticisme à un fournisseur infonuagique

L'agnosticisme à un fournisseur infonuagique d'un système peut être recherché pour diverses raisons, comme pour simplifier le développement de composants logiciels pour plusieurs plateformes (e.g. qui doivent être exécutés aussi bien sur site que sur divers fournisseurs infonuagiques), ou bien pour simplifier les migrations depuis un nuage-hôte si les conditions d'hébergement venaient à changer et ne plus être compatibles avec les exigences du système (e.g. tarifs, politiques).

Une manifestation de cette exigence est la volonté d'utiliser des logiciels libres.

Ainsi, dans le cadre d'une migration d'applications relevant de la sûreté de l'État, l'un de nos interlocuteurs a fait savoir la nécessité d'utiliser des logiciels libres afin d'éviter tout enfermement propriétaire, que cela concerne le fournisseur infonuagique comme la société prestataire qui administre le système. La crainte était

d'être, in fine, prisonnier des tarifs d'une entreprise. Pour cette raison, un service de *Kubernetes* à la demande a été choisi pour héberger leurs portails web plutôt qu'un *PaaS* propriétaire : « *On a choisi Kubernetes, car c'est un standard industriel libre [...] pour que l'entreprise soit le plus indépendant possible vis-à-vis d'autres entreprises* ».

6. Propositions de classification des exigences

Suite à la mise en évidence de ces dix exigences de haut niveau, nous avons cherché à établir une classification. Classiquement, l'ingénierie des exigences distingue les exigences fonctionnelles, qui décrivent les fonctions du système à développer, et non-fonctionnelles, qui décrivent les qualités désirées du système (Kotonya, Sommerville, 1998). En général, une migration ne transforme pas le produit, mais son support. Dès lors, la plupart des exigences que nous avons identifiées touchent peu au fonctionnel, cette classification n'est pas ici pertinente.

De leur côté, les approches orientées buts distinguent les *buts durs* qui peuvent être satisfaits sans ambiguïté, et des *buts mous* (*soft-goal*), ou *qualités*, qui décrivent une direction sans critère de satisfaction clair (Dalpiaz *et al.*, 2016 ; Lamsweerde, 2001). D'après les entretiens menés, certaines exigences peuvent être classées aisément dans l'une de ces deux catégories. Par exemple, l'exigence de souveraineté numérique est un *but dur*, car elle fixe sans ambiguïté ce qui est acceptable pour la législation que suit un fournisseur infonuagique ou la localisation d'une donnée. De même, si tous nos interlocuteurs ont manifesté le besoin de minimiser la maintenance opérationnelle de leur système, aucun chiffre n'a été avancé ; cette exigence est donc un *but mou*. Cependant, certaines exigences, telles qu'énoncées dans les entretiens, peuvent être rangées à la fois comme un *but mou* ou *dur*. Par exemple, minimiser le temps de migration est, selon le contexte, un *but mou*, si l'on cherche seulement à minimiser le temps, ou un *but dur*, si la migration doit être terminée avant une date fixée. Par ailleurs, même si ce n'est pas apparu lors de nos entretiens, les exigences comme la minimisation des coûts ou l'amélioration des performances sont propices à être mesurées (e.g. "diminuer la latence réseau" contre "la latence réseau ne dépasse pas tel seuil") et donc leur classification dépendrait du contexte également. Pour ces raisons, nous ne considérerons pas cette classification.

Finalement, notre choix s'est porté sur la dichotomie : développement logiciel et exploitation du système (DevOps). Il nous faut cependant prendre en compte un troisième élément : le fournisseur infonuagique. En effet, de nombreuses exigences reposent sur la capacité du fournisseur infonuagique à délivrer les services adéquats. Par conséquent, nous proposons la classification suivante :

- Les exigences liées au **développement logiciel**. Les choix de services découlant de ces exigences vont interférer avec le développement (dans le cadre de la migration, il s'agit d'adaptations) des logiciels. Les exigences liées au développement logiciel sont : **maximiser l'agnosticisme à un fournisseur infonuagique, minimiser le besoin de nouvelles expertises, minimiser les coûts et la durée de migration**.

- Les exigences liées à l'**exploitation du système**. Les efforts investis dans l'exploitation du système dépendent des choix de services découlant de ces exigences. Les exigences liées à l'exploitation du système sont : **maximiser la fiabilité du système, minimiser le besoin de nouvelles expertises, minimiser la maintenance opérationnelle**.
- Les exigences liées au **fournisseur infonuagique**. Ces exigences sont satisfaites uniquement par des propriétés du fournisseur. Les exigences liées au fournisseur infonuagique sont : **garantir la souveraineté numérique, maximiser les performances, minimiser les coûts de l'environnement, maximiser la sécurité**.

7. Travaux futurs

Sur la base de l'analyse présentée dans cet article, nous avons commencé une enquête que nous diffusons plus largement auprès d'experts industriels. Nous attendons de cette enquête qu'elle nous permette de confirmer la pertinence des exigences que nous avons présentées ici et d'identifier de nouvelles exigences de haut niveau qui n'auraient pas été évoquées dans nos entretiens comme, par exemple, les exigences relatives aux impacts environnementaux des systèmes infonuagiques, qui sont pourtant de plus en plus souvent imposées dans les appels d'offres, en particulier pour les marchés publics.

Notre classification n'est pas la seule possible. Classiquement, dans la gestion de projets, on identifie trois principales contraintes : la maximisation de la qualité, la minimisation des délais et la minimisation des coûts. Il est peut-être possible de classer les exigences que nous présentons dans cet article comme des raffinements de ces contraintes. En effet, certaines exigences visent à minimiser les coûts (minimiser les coûts de migration, minimiser le besoin de nouvelles expertises, minimiser la maintenance opérationnelle), d'autres à minimiser les délais (minimiser la durée de migration, minimiser le besoin de nouvelles expertises) et enfin certaines à maximiser la qualité du système (maximiser la sécurité, maximiser l'agnosticisme à un fournisseur infonuagique, garantir la souveraineté numérique, maximiser les performances, maximiser la fiabilité du système). Des travaux futurs devraient explorer les classifications possibles et chercher les plus efficaces pour guider la gestion des projets de migration.

Nos dix exigences ne sont pas indépendantes. Par exemple, minimiser la maintenance opérationnelle implique des développements pour utiliser des services hautement infogérés et rentre donc en conflit avec l'exigence de minimiser la durée de migration. Afin d'assister la prise de décision dans le cadre d'un compromis entre les contraintes d'un projet de migration, il serait intéressant d'identifier les liens entre nos exigences de haut niveau. Une fois ces liens identifiés, il pourrait être intéressant de définir un cadre méthodologique, basé sur les approches orientées buts, pour déterminer les compromis acceptables, et ainsi mieux appréhender la complexité cognitive inhérente à une migration vers un nuage public.

Les *stratégies de migration* classifient les migrations selon les activités et les choix de conception de l'environnement. Si la littérature scientifique et industrielle propose plusieurs catégorisations de stratégie de migration, il manque une méthode holistique pour déterminer quelle stratégie employer dans une situation donnée. Nous retrouvons ce constat dans les travaux de Zhao et Zhou (2014) qui identifient des défis à traiter sur la thématique des stratégies de migration. Une telle méthode holistique devrait prendre en compte les exigences que nous avons identifiées ici.

Enfin, l'évaluation de la satisfaction de certaines de ces exigences dès la conception n'est pas triviale. C'est par exemple le cas de l'exigence "minimiser les coûts de l'environnement" qui requiert d'estimer le coût de l'environnement infonuagique, qui dépend de la charge utilisateur à tout instant. Dès lors, il est coûteux pour une organisation d'identifier les situations dans lesquelles il est nécessaire de faire des compromis, et de formuler des compromis acceptables quand c'est nécessaire. En outre, des travaux apportent des approches pour évaluer individuellement certaines qualités d'un environnement infonuagique, comme ceux de Gesvindr *et al.* (2017) pour la performance des *PaaS* ou Belli *et al.* (2016) pour l'estimation des coûts des *IaaS*, mais il manque d'approches permettant de comparer entre elles des architectures proposant les mêmes fonctionnalités tout en étant très différentes (e.g. une architecture basée sur des *IaaS* et une autre *serverless*).

8. Conclusion

La migration d'un système vers un nuage permet aux entreprises de bénéficier des avantages de l'informatique en nuage. Si la littérature décrit plusieurs méthodes pour satisfaire certaines exigences, elle fait fi des contraintes que ces dernières s'imposent entre elles. En outre, il manque de travaux sur l'identification des exigences que peut avoir l'industrie.

Dans cet article, nous avons présenté l'analyse de retours de professionnels de l'informatique en nuage au sujet de la migration de systèmes vers un nuage. Ces retours ont été collectés à travers une série d'entretiens semi-formels. Cette analyse a mis en évidence dix exigences de haut niveau qui dirigent le choix des services infonuagiques et de leur configuration lors de la conception, pour lesquelles nous avons proposé une classification.

Bibliographie

- Amiri M., Khanli L. M. (2017). Survey on prediction models of applications for resources provisioning in cloud. *J. Netw. Comput. Appl.*, vol. 82, p. 93-113.
- Belli O., Loomis C., Abdennadher N. (2016). Towards a cost-optimized cloud application placement tool. *IEEE International Conference on Cloud Computing Technology and Science*, p. 43-50.
- Dalpiazz F., Franch X., Horkoff J. (2016). iStar 2.0 langage guide. *CoRR*, vol. abs/1605.07767.

- DiCicco-Bloom B., Crabtree B. F. (2006). The qualitative research interview. *Medical education*, vol. 40, n°4, p. 314-321.
- Ferry N., Rossini A., Chauvel F., Morin B., Solberg A. (2013). Towards model-driven provisioning, deployment, monitoring, and adaptation of multi-cloud systems. *2013 IEEE Sixth International Conference on Cloud Computing*, p. 887-894.
- Gesvindr D., Buhnova B., Gasior O. (2017). Quality evaluation of PaaS Cloud application design using generated prototypes. *IEEE International Conference on Software Architecture*, ICSA 2017, p. 31-40.
- Kotonya G., Sommerville I. (1998). *Requirements engineering – processes and techniques*. John Wiley & Sons.
- Lamsweerde A. van (2001). Goal-oriented requirements engineering: A guided tour. *5th IEEE International Symposium on requirements engineering*, p. 249.
- McCracken G. (1988). *The long interview*. SAGE Publications, Inc.
- Mell P., Grance T. (2011). *The NIST definition of Cloud Computing*. Rapport technique n°800-145.
- Quinton C. (2014). *Cloud Environment Selection and Configuration: a software product lines-based approach*. Thèse de doctorat, Université Lille I.
- Zhang M., Ranjan R., Nepal S., Menzel M., Haller A. (2012). A declarative recommender system for cloud infrastructure services selection. *Economics of grids, clouds, systems, and services – 9th International Conference*, vol. 7714, p. 102-113.
- Zhao J., Zhou J. (2014). Strategies and Methods for cloud migration. *Int. J. Autom. Comput.*, vol. 11, p. 143-152.